

© 2011 by Sergey Popov. All rights reserved.

THREE ESSAYS IN ECONOMICS

BY

SERGEY POPOV

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Mark Daniel Bernhardt, Chair and Director of Research
Professor In-Koo Cho
Associate Professor George Deltas
Professor Mattias Polborn
Lecturer Rui Zhao

Abstract

This is a study of market organization in very different settings.

In the first chapter, I study how the choices by students to “rush” fraternities, and those of fraternities of whom to admit, interact with the signals that firms receive about student productivities to determine labor market outcomes. Students understand that fraternities value both academic ability and “socializing values”, and will screen out students with lesser academic abilities unless they have especially high socializing values. Students also understand that firms will see whether or not they are in a fraternity, and will use that information in their hiring decisions. I characterize possible equilibrium outcomes and then estimate the model. I identify sufficient conditions for highly-able students to perceive participation in a fraternity as tainting, so that students of intermediate abilities are over-represented in fraternity membership. That is, in equilibrium most less able students apply to fraternities, but are rejected unless they have high socializing values, while most able students do not apply to avoid being tainted in labor market outcomes from being mixed in with less able fraternity members. I show that my theory can reconcile the data on cumulative GPAs of fraternity and non-fraternity members at the University of Illinois. I structurally estimate a three signal version of the model and show the welfare and wage impacts of the fraternity for different types of students.

In the second chapter, I study the sovereign borrowing market. Sovereign default often affects country’s trade relations. The defaulter’s currency depreciates while trade volume falls drastically. To explain this connection, I propose a model to incorporate real depreciation along with sovereign bankruptcy. Defaulters must exchange more of their own goods for imports, which stimulates an adjustment to the equilibrium exchange rate. I demonstrate that a default episode can imply up to a 30% real depreciation. This matches the depreciations observed in crisis events for developing countries. To avoid this, countries are willing to maintain borrowing obligations up to a realistic level of debt.

In the third chapter, I develop a model of strategic grade determination by universities distinguished by their distributions of student academic abilities. Universities choose grading standards to maximize the total wages of graduates, taking into account how the grading standards affect firms’ productivity

assessment and job placement. Job placement and wages hinge on a firm's productivity assessment given a student's university, grade and productivity signal. I identify conditions under which better universities set lower grading standards than is socially optimal, exploiting the fact that firms cannot distinguish between "good" and "bad" "A"s, while worse universities set excessively strict grading standards. In particular, a social planner finds it optimal to set stricter grading standards at better universities, whereas the opposite happens in equilibrium. I then show how increases in skilled jobs drive grade inflation, and determine when grading standards fall faster at better universities. These findings are consistent with the evolution of average cumulative GPAs at different schools over the past 50 years.

Acknowledgments

This project would not have been possible without the support of many people. Many thanks to my adviser, Dan Bernhardt, who read my numerous revisions and helped to formulate my thoughts consistently. Also thanks to my committee members, George Deltas, Mattias Polborn, In-Koo Cho, and Rui Zhao, who offered guidance and support. And finally, thanks to my classmates and numerous friends who endured this long process with me, always offering support and love.

Table of Contents

List of Tables	vii
List of Figures	viii
Chapter 1 Fraternities and Labor Market Outcomes	1
1.1 Introduction	1
1.1.1 Related Literature	3
1.1.2 Fraternities and Sororities	5
1.2 The Fraternity Game	5
1.2.1 Student's Problem	10
1.2.2 Fraternity's Problem	11
1.2.3 Existence of equilibrium	12
1.3 Three Signal Economy	13
1.4 Single-peaked Equilibria	16
1.5 Empirical Analysis	20
1.6 Conclusion	29
Chapter 2 Equilibrium Sovereign Default with Endogenous Exchange Rate Depreciation 30	30
2.1 Introduction	30
2.1.1 Patterns of Trade in Default	31
2.1.2 Trade Channel Penalty	32
2.2 The Model	34
2.2.1 Timing	35
2.2.2 Representative Household and Government	35
2.2.3 Importers	36
2.2.4 International Financial Markets	37
2.2.5 Recursive Equilibrium	37
2.3 Quantitative Evaluations	39
2.3.1 Quarterly Data	39
2.3.2 Solving for the Recursive Equilibrium	40
2.3.3 Welfare Comparisons	41
2.3.4 Equilibrium Results	42
2.3.5 The Effect Of Penalty Levels	44
2.3.6 Comparative Statics	48
2.4 Conclusion	50
Chapter 3 University Competition, Grading Standards and Grade Inflation	51
3.1 Introduction	51
3.2 The Competition Between Universities	54
3.2.1 Equilibrium	57
3.3 No Heterogeneity in Social Skills	58
3.4 Heterogeneous Social Skills	62
3.5 Conclusion	65

Appendix A	Fraternities and Labor Market Outcomes	67
Appendix B	Equilibrium Sovereign Default with Endogenous Exchange Rate Depreciation	72
B.1	Centralized Equilibrium	72
B.2	Estimation	73
B.2.1	Consumer's Utility Parameters	74
B.2.2	Import-Export Equation	74
B.3	Solving for First Best	76
Appendix C	University Competition, Grading Standards and Grade Inflation	77
References		81

List of Tables

1.1	Estimates	24
1.2	Estimates for Gender-Specific Models	26
1.3	Estimates for Gender-Specific Models: Likelihood-Based	28
2.1	One year effect of sovereign default.	31
2.2	Parameter values, quarterly data.	40
2.3	Welfare comparison.	41
3.1	Evolution of GPAs at selected universities.	52
B.1	Estimation results, output dynamics, One star denotes 1% significance.	73
B.2	Estimation results, output dynamics.	74
B.3	Estimation results, import-export conditions.	75

List of Figures

1.1	Equilibrium with Completely Informative Signals.	9
1.2	Equilibrium with Completely Uninformative Signals.	10
1.3	Unconstrained admissions set B	12
1.4	Application-unconstrained equilibrium.	15
1.5	Application constrained fraternity game equilibrium.	16
1.6	Equilibrium Wages. The solid line is the expected wage of a fraternity member (as a function of θ), and the dashed line is the expected wage of a nonmember.	17
1.7	Application-constrained equilibrium for different capacities; arrows denote the intersection of fraternity's and students' choices in corresponding equilibrium.	18
1.8	Conditional Distributions of GPAs	21
1.9	Estimated Unconstrained Model and Structural Model Fits	24
1.10	Impact of Fraternity on Student Welfare	28
1.11	Impact of Wage Redistribution on Student Welfare	28
2.1	Dynamics of exchange rates in Paraguay and Ukraine, nominal and real. Exchange rates are normalized so that maximum value on sample is 1.	32
2.2	The export-import transformation function and the effect of the trade channel penalty	38
2.3	Borrowing policy. Curves stop on right at the point of default.	42
2.4	Current account conditional on output in stationary distribution	43
2.5	Exchange rate conditional on output in stationary distribution	43
2.6	Default decision borrowing threshold	44
2.7	Exchange rate drop due to default	45
2.8	Domestic good consumption (left) and foreign good expenditure (right)	46
2.9	Trade balance (left) and capital account (right) changes at default.	46
2.10	Comparative statics.	47
3.1	Comparison of Equilibrium and Social Planner outcomes.	60
3.2	Equilibrium and Social Planner outcomes.	63
B.1	Time series of $u_t = \frac{\Delta m_t}{\Delta x_t} \frac{x_t}{m_t}$, "sample elasticity."	75

Chapter 1

Fraternities and Labor Market Outcomes

1.1 Introduction

To many, the word “fraternities” brings to mind images of beer, parties and fun. Yet, fraternity membership also enters prominently the job seeking process of many students: resumes often devote scarce space to highlighting a student’s society memberships in addition to the standard information about education, work experience, awards, etc. This suggests that fraternity membership helps employers evaluate a person’s productivity. On first impression it is not clear why fraternity or sorority membership should matter for labor market outcomes. In particular, while fraternities make significant time demands — members must spend considerable time picking up trash on highways, raising money for charitable causes, and so on — these activities appear largely unrelated to skill development for future careers. Nonetheless, fraternities draw many applicants who eagerly spend money and devote time to these activities, and employers seem to weigh membership information positively.

We develop a theory of fraternity membership and filtering by firms that makes sense of these observations. Students are distinguished by a fraternity socializing value and their productivity as a worker. Fraternities value both the future wages generated by members and their socializing values. Firms combine information in noisy signals about student productivities with fraternity membership status to set wages. To emphasize the key economic forces, we suppose that fraternity socializing values are not directly valued by firms and also that they are uncorrelated with worker productivities. We further assume away all standard club features for fraternities as in Buchanan (1965), so that there are no consumption spillovers due to the presence of other students. So, too, we assume away any networking services that a fraternity might provide. As a result, the fraternity membership statuses of other students only affect job market outcomes for a given student via the equilibrium beliefs that firms form about the distribution of abilities of fraternity members and non-members.

We first identify sufficient conditions for fraternity membership not to matter for job market outcomes. In particular, we show that if the signal that firms receive about a student’s productivity is either perfectly informative or perfectly noisy, then equilibrium wages do not depend on a student’s fraternity membership.

As a result, whether a student rushes a fraternity depends only on his fraternity socializing value. If signals are perfectly informative, fraternities trade off between productivity and socializing value in admission, but a student's wage will equal his known productivity, rendering membership irrelevant for labor market outcomes. If, instead, signals are perfectly noisy, a fraternity would like to commit to excluding low ability students with high socializing value, and to accepting high ability students with low socializing value. However, with perfectly noisy productivity signals, firms have no source other than fraternity membership for evaluating a student's ability, so that fraternities weigh only socializing values in admission. As a result, fraternity membership conveys no information to firms about ability, so that wages do not hinge on fraternity membership.

In sum, we show that for fraternity membership to affect job market outcomes, firms must receive signals about a student productivities that are noisy, but not perfectly so. Then, because more productive students tend to earn higher wages, fraternities trade off between productivity and fraternity socializing values when deciding which pledge applicants to accept. In particular, fraternities accept students with low socializing values who are sufficiently able. Students may face a different type of trade off—more able students may incur a labor market cost from joining a fraternity, as their fraternity membership may lump them in with intermediate quality students, making it harder for the able students to distinguish themselves in the eyes of firms. In such a situation, sufficiently more able students may be reluctant to pledge fraternities.

We then turn to a three-signal setting in which we can explicitly solve for the multiple equilibria that emerge in the fraternity game. We identify three types of equilibria: (a) an “empty fraternity” equilibrium in which no student applies to the fraternity, supported by beliefs of firms that any student who joins the fraternity is especially lacking in ability; (b) a “single-peaked” equilibrium in which most fraternity members have intermediate abilities—less able students apply, but are rejected unless they have high fraternity socializing values, while very able students who do not have very high fraternity socializing values do not apply; and (c) an equilibrium in which employer beliefs about the abilities of fraternity members are more optimistic—so that fraternity membership would increase the expected wage of each student type. In these latter two equilibria, relatively low ability students expect higher wages if they gain fraternity membership than if they do not; while in the second equilibrium, but not the third, higher ability students may anticipate lower wages if they join. That is, fraternity membership may taint labor market outcomes for high ability students, but not low.

We return to a more general signal framework, in which we only assume that the conditional distribution of ability signals that firms receive satisfies the monotone likelihood ratio property—more able students are more likely to generate higher signals. We then provide gross sufficient conditions for non-trivial equilibria to have the single-peaked feature. In particular, these sufficient conditions imply that the wage premium

due to fraternity membership declines with ability—the lowest ability fraternity members always receive a particularly large wage premium, as their membership ensures that they are separated away from all lower ability types, and are mixed in with relatively higher proportions of more able types, while high ability types gain less (or lose) from being mixed in with less able fraternity members. The single-peaked equilibrium then emerges, due to the filtering by fraternities of low ability students, and, when membership costs are of an appropriate magnitude relative to socializing values, the reluctance of high ability students to join.

Finally, we investigate whether equilibria of our three-signal model are consistent with actual practice. To do this, we obtain data on the distributions of cumulative GPAs of seniors at the University of Illinois for fraternity members and non-members. We find that the percentage of students with a given GPA who are fraternity members is a sharply single-peaked function of GPA. This relationship indicates that the data are inconsistent with both pure signaling and pure screening theories, as well as with basic networking theories in which ability and networking are either substitutes or complements in the wage-generation process. Such explanations predict a monotone relationship between GPA and the fraction of students with that GPA who are fraternity members.

We then show that even with a simple three signal structure, the equilibrium of our model in which most fraternity members have intermediate abilities—where high ability students are tainted by membership in fraternities—can generate a distribution over the probability of membership conditional on ability that closely mirrors the distribution over the probability of fraternity membership conditional on GPA found in the data. We back out plausible estimates of primitives—the time costs of fraternity participation, and the tradeoffs of both students and fraternity between socializing values and future wages. We use these estimates to derive how the presence of the fraternity affects the welfare of different student types.

We next review the literature and then provide a brief overview of fraternities. In section 2 we develop our model and analysis. Section 3 considers a three signal setting. Section 4 returns to a general setting and provides sufficient conditions for all non-trivial equilibria to exhibit the hump-shaped pattern that we find in the data. Section 5 provides our empirical analysis of fraternities. Section 6 concludes. Some proofs are in an appendix.

1.1.1 Related Literature

Our model and analysis is closest to the endogenous statistical discrimination literature, which focuses on the costly signaling actions that privately-informed individuals may take to influence the hiring decisions of firms. Moro and Norman (2004) consider a setting in which individuals choose whether to make a costly investment

in education, when firms receive noisy signals of that investment.¹ They show how a productivity-irrelevant aspect such as racial identity can affect investment choices if firms believe that one population is more likely to invest. In particular, they show that multiple equilibria can exist, one where beliefs do not depend on race, and one where they do.² Austen-Smith and Fryer (2005) consider a setting in which a “peer group” generates additional utility from leisure for its members, and this drives members to shift time allocation toward less education, which leads to lower wages. Those who are rejected by the peer group study more, “acting white”.

In contrast to this literature, our environment features three active strategic players—students, fraternity and firms—and rather than information being conveyed to firms via signaling by students, it is conveyed by equilibrium outcome of the application/acceptance game between students and the fraternity. Thus, the information content of fraternity membership is determined by the equilibrium of the game played between the fraternity and different student types, where both the students and the fraternity take into account that the fraternity membership outcomes are observed by a third party, the firm. Most crucially, this means that ours is **not** a signaling model in the classical sense of Spence (1973). To see this, recognize that if the fraternity were not to weigh ability in its admissions, then fraternity membership would not convey any information about ability, and hence not affect wages. Moreover, one of our central propositions shows that when the fraternity does weigh ability in admissions and admits sufficiently few students, then conditional on a given socializing value, the net benefit of fraternity membership is **monotonically decreasing** in ability. Even though, in equilibrium, the unconditional wage of fraternity members exceeds those of non-members, the standard single-crossing property of signaling could not be more violated. The single-peak nature of the equilibrium in our economy is not just a consequence of the fact that firms receive two pieces of information (fraternity membership and exogenous signals of ability), but rather it is due to the filtering out by fraternities of most low ability applicants (who apply in large numbers), combined with the increasing reluctance of higher ability students to apply. Thus, despite the added analytical challenges associated with information transmission to firms being generated by a game between two active agents, we show how this strategic interaction matters, and generate sharp predictions about equilibrium behavior.

Our economy also differs from this signaling literature in several other key aspects. Fraternity membership is endogenous and observed, but is not productivity-enhancing; while in the statistical discrimination literature, race is exogenous, and investment is unobserved, but productivity-enhancing. Moreover, in contrast to the signaling literature, in our model for fraternity membership to influence labor market outcomes, firms must also have access to additional information about a worker’s ability.

¹See Coate and Loury (1993), Fang and Norman (2006) and Norman (2003) for related models.

²Mailath, Samuelson and Shaked (2000) develop a related search model in which firms choose which populations to search, and each population makes investments in skills based on beliefs about firm search intensities. Again, asymmetric search intensities can give rise to asymmetric investment choices in two otherwise identical populations.

1.1.2 Fraternities and Sororities

To ease presentation, we drop gender differences and refer to both fraternities and sororities as “fraternities”. The first club-like fraternity with a centralized organization was Kappa Alpha Society, founded in 1825. For a history and current status of fraternities, see Jack L. Anson and Robert F. Marchesani (1991).

Fraternities require pledge applicants to submit extensive information about themselves: their school GPA, recommendations, interests and useful skills. Fraternities devote far more time to evaluating applicants than do potential employers. In particular, almost all fraternity applicants are interviewed, and applicants take part in an extensive series of activities during the evaluation process. For example, Sigma Chi requires a potential member to spend one year working for the fraternity before the pledge. This suggests that fraternities are well-situated to evaluate a pledge applicant’s ability, so that fraternity membership can provide firms with valuable information.

Fraternities rely on membership fees and donations to fund activities. A substantial share of a fraternity’s income comes from alumni donations. Because high income alumni donate more, fraternities care about the future job market outcomes of members. An indication of the value that fraternities place on productive members is that GPA-based stipends to fraternity members are widespread. Fraternities frequently reject pledge applicants. Conversely, many highly-productive students choose not to apply to fraternities. Finally, students almost never join more than one fraternity. This reflects both secrecy issues (secret handshakes, for example, allow one member to verify the membership status of others), and because fraternity activities are quite time-consuming.

1.2 The Fraternity Game

Students. There is a measure 1 of students. A student is fully described by his future employment productivity θ and his fraternity socializing value, μ . Students have separable preferences over income and fraternity membership: a non-member’s payoff corresponds to his expected net lifetime income, M , and a fraternity member with socializing value μ derives utility

$$M + n\mu,$$

where $n > 0$ is the weight that students place on socializing values. M equals the student’s expected lifetime future wage minus the monetary value c of the time costs of fraternity service activities. Note that to sever all links with the club-good literature, we assume away any externalities from the socializing values of other

fraternity members. So, too, to ensure that there is no direct link between productivity and membership, we assume that θ and μ are uncorrelated in the population. That is, the density over θ and μ is

$$h(\theta, \mu) = h_\theta(\theta)h_\mu(\mu),$$

where the bounded supports of θ and μ are given by $[\underline{\theta}, \bar{\theta}]$ and $[\underline{\mu}, \bar{\mu}]$, respectively, and $\underline{\theta}$ and $\bar{\mu}$ are both positive. The associated cdf is H , and the measure m , used in some proofs, is based on H . We emphasize that while we believe that socializing skills and productivity may be correlated in practice, we assume such correlation away in order to highlight the impact of application decisions by students and filtering by fraternities on the equilibrium distribution of abilities in the fraternity.³

Fraternity. There is a single representative fraternity that chooses which “rush” applicants to admit. The fraternity cares about both the future market wages that its members will obtain, and the socializing values of its members. For simplicity, we assume that the fraternity has separable linear preferences over wages and socializing values, so that the fraternity’s payoff from members (θ, μ) in the set C of fraternity members is given by

$$\int_{(\theta, \mu) \in C} [W_1 E_{\tilde{\theta}}(w_C(\tilde{\theta}|\theta)) + W_2 \mu] h(\theta, \mu) d\theta d\mu,$$

where $W_1 > 0$ and $W_2 \geq 0$ are the weights that the fraternity places on wages and socializing values of its members.

We assume that the fraternity is limited by space constraints to admitting at most a measure Γ of students: in practice, a fraternity house has a limited number of bedrooms. This means that the fraternity trades off between μ and θ in admission—trading off future higher contributions from more able and hence wealthier alumni against their social contribution.

Our analysis is qualitatively unaffected by alternative model formulations that preserve the fraternity’s incentive to trade-off between socializing values and ability in admissions. For example, qualitatively identical outcomes emerge if the fraternity did not face a space constraint, but instead cared about the average socializing value of its members (say due to externalities), in addition to the future wages that members earn. So, too, outcomes are qualitatively unaffected if the fraternity, rather than facing a space constraint, incurred costs that were a convex function of the measure of members (say due to cramming more students into each room), or if the fraternity cared about the market value of the time contributions of its members.

Firms. After graduation, students are employed by firms. We assume that several risk-neutral firms make

³Obviously, if social skills and productivity are positively correlated in the population, and the fraternity values social skills, then this exogenous correlation will lead to fraternity members receiving higher wages than non-members; we wanted to avoid building this result trivially into our model. We assume away any network services that a fraternity might provide for the same reason.

simultaneous wage offers to students. The firms do not observe an individual student's productivity θ or fraternity socializing value μ . However, firms do observe whether a student is a member of a fraternity. Firms also observe a common signal $\tilde{\theta}$ about the student's productivity θ , where $\tilde{\theta}$ is distributed according to $F_{\tilde{\theta}}(\cdot|\theta)$. More able students are more likely to generate higher signals: $F_{\tilde{\theta}}(\tilde{\theta}|\theta)$ is strictly decreasing in θ for all $(\theta, \tilde{\theta})$ with $F_{\tilde{\theta}}(\tilde{\theta}|\theta) \in (0, 1)$. Competition drives firms to offer each individual a wage equal to his expected productivity given his fraternity membership status and ability signal, $\tilde{\theta}$.

There are four stages to our "fraternity rush" game. At stage one, each student type (θ, μ) decides whether to apply for fraternity membership. We let $a(\theta, \mu)$ be an indicator function taking on the value 1 if student type (θ, μ) applies, and taking on the value 0 if the student type does not apply. We sometimes use the set $A = \{(\theta, \mu) | a(\theta, \mu) = 1\}$. At stage 2, the fraternity chooses which applicants to accept. We let $b_A(\theta, \mu)$ be an indicator function taking on the value 1 if, given the set of applicants A , the fraternity would admit a student type (θ, μ) who applied, and taking on the value 0 otherwise. We use $B_A = \{(\theta, \mu) | b_A(\theta, \mu) = 1\}$ to represent the set of admitted student types. Then, the set of fraternity member types is $C_A = \{(\theta, \mu) | a(\theta, \mu)b_A(\theta, \mu) = 1\}$, and the set of nonmembers is $\bar{C}_A = \{(\theta, \mu) | a(\theta, \mu)b_A(\theta, \mu) = 0\}$. At stage 3, firms see whether an individual is a fraternity member, and they see a noisy signal of his ability, but do not observe their types—firms must form beliefs about which student types actually join the fraternity. Let $\rho_F(\theta, \mu)$ denote firm beliefs about fraternity membership for each type (θ, μ) , where $\rho_F(\theta, \mu) = 1$ if firms believe type (θ, μ) is a member of the fraternity, and $\rho_F(\theta, \mu) = 0$ if not. Finally, $w_C(\tilde{\theta})$ denotes the wage of a fraternity member who emits the signal $\tilde{\theta}$, and $w_{\bar{C}}(\tilde{\theta})$ denotes the wage of a non-member who generates signal $\tilde{\theta}$. At stage 4, a worker with productivity θ produces output with value θ .

Equilibrium. An equilibrium is a collection of functions, $\{a^*(\theta, \mu), b_A^*(\theta, \mu), w_C^*(\tilde{\theta}), w_{\bar{C}}^*(\tilde{\theta})\}$ and firm beliefs $\rho_F^*(\theta, \mu)$ such that

i) Students optimize: $a^*(\theta, \mu) = 1$ if $E[w_C^*(\tilde{\theta})|\theta] + n\mu - c \geq E[w_{\bar{C}}^*(\tilde{\theta})|\theta]$; 0 otherwise. We let A^* be the associated set of fraternity applicants.

ii) For every A the fraternity optimizes: B_A^* solves

Problem 1

$$\max_{B_A} \int_{(\theta, \mu) \in A \cap B_A} [W_1 E[w_C^*(\tilde{\theta})|\theta] + W_2 \mu] h(\theta, \mu) d\theta d\mu \quad (1.1)$$

subject to $m(A \cap B_A) \leq \Gamma$.

iii) wages are competitive given beliefs by firms $\rho_F^*(\theta, \mu)$:

$$w_C^*(\tilde{\theta}) = \frac{\int_C \theta h(\theta, \mu) \rho_F^*(\theta, \mu) f_{\tilde{\theta}}(\tilde{\theta}|\theta) d\theta d\mu}{\int_C h(\theta, \mu) \rho_F^*(\theta, \mu) f_{\tilde{\theta}}(\tilde{\theta}|\theta) d\theta d\mu}; \quad w_{\bar{C}}^*(\tilde{\theta}) = \frac{\int_{\bar{C}} \theta h(\theta, \mu) (1 - \rho_F^*(\theta, \mu)) f_{\tilde{\theta}}(\tilde{\theta}|\theta) d\theta d\mu}{\int_{\bar{C}} h(\theta, \mu) (1 - \rho_F^*(\theta, \mu)) f_{\tilde{\theta}}(\tilde{\theta}|\theta) d\theta d\mu}.$$

iv) Firm beliefs are consistent with choices of student types and fraternity: For a.e. (θ, μ) , $\rho_F^*(\theta, \mu) = a^*(\theta, \mu)b_A^*(\theta, \mu)$.

Off-equilibrium path characterizations are not intrinsically interesting, and to ease presentation, we only characterize equilibrium path outcomes. Moreover, measure zero perturbations of the fraternity's acceptance set B_{A^*} are uninteresting—any measure zero perturbation to B_{A^*} is also part of an equilibrium, so we focus on a best response of the fraternity that is a **good set**, i.e., a set B_{A^*} that is equal to the closure of its own interior. Also observe that it is almost immediate that equilibria to our single fraternity game are also equilibria to a multi-fraternity game in which the ability/socializing value composition of applicants to each fraternity is the same, and hence the fraternities tradeoff in the same way between μ and θ in admission.⁴ Finally, to simplify notation, we omit the A^* index on the equilibrium acceptance set.

We begin by providing conditions under which fraternity membership has no effect on labor market outcomes.

Proposition 1 *Suppose that firms either receive perfect signals about students, i.e., $\tilde{\theta} = \theta$, a.e., θ , or firms receive perfectly uninformative signals, $F_{\tilde{\theta}}(\cdot|\theta) = F_{\tilde{\theta}}(\cdot|\theta')$, for all θ, θ' . Then, in equilibrium, a student's wage does not depend on whether he is in the fraternity or not, i.e., $w_C(\tilde{\theta}) = w_{\bar{C}}(\tilde{\theta})$, for all $\tilde{\theta}$. Hence, a student type (θ, μ) applies for membership in the fraternity if and only if $n\mu - c \geq 0$.*

Suppose signaling is perfect. Then $w_C(\tilde{\theta}) = w_{\bar{C}}(\tilde{\theta}) = \tilde{\theta} = \theta$ (a.e.). Optimization by students then implies that a student type (θ, μ) applies if and only if $n\mu - c \geq 0$, independently of θ . In contrast to students, the fraternity selectively admits higher θ applicants who will earn higher wages. In particular, the fraternity trades off between μ and θ in admission; letting $\mu_B(\theta)$ denote the boundary of the admission set, indifference implies that the boundary has slope $\frac{d\mu_B(\theta)}{d\theta} = \frac{-W_1}{W_2}$. Figure 1.1 illustrates the equilibrium. The solid line is the fraternity's equilibrium cutoff rule—all types to the right of the line who apply are accepted, while all those to the left are rejected. The vertical dashed line represents the accept-or-reject line of students—student types to the right apply in equilibrium, i.e., are in the set A . Hence, the equilibrium set C of fraternity members consists of those types to the right of both the dashed and solid lines, and the measure of the set C is at most Γ .

If, instead, signals are completely uninformative, then all individuals receive the same wage, $w_C(\tilde{\theta}) = w_{\bar{C}}(\tilde{\theta}) = E[\theta]$. A fraternity would like to commit to excluding low θ students with high socializing values μ , and to accepting high θ students with low socializing values. However, since firms have no source other than fraternity membership for evaluating a student's ability, all fraternity members must receive the same

⁴For example, with 2 fraternities, each with capacity $\frac{\Gamma}{2}$, it is an equilibrium for each fraternity to set the same acceptance set, and for students who will be admitted to fraternities in equilibrium, to divide their applications between fraternities in such a way that the distributions of ability and socializing values of accepted students are the same at both fraternities, and for firms to set wages consistently with such behavior.

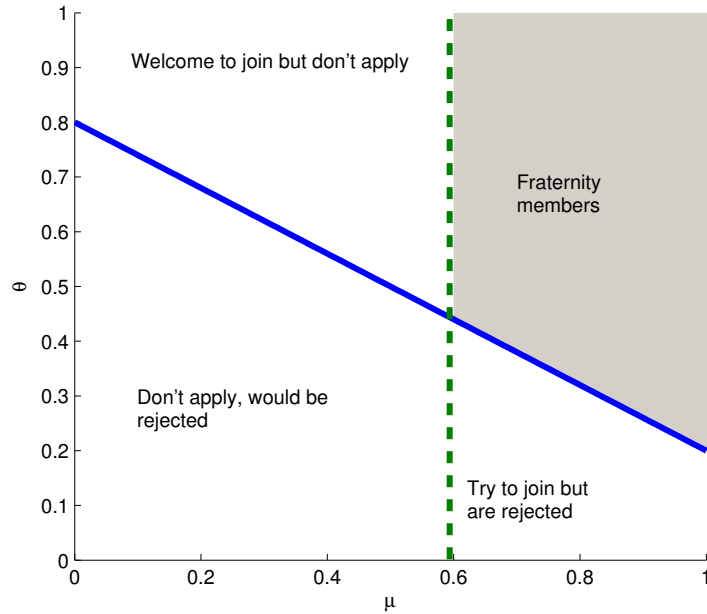


Figure 1.1: Equilibrium with Completely Informative Signals.

wage. But then, given *any* beliefs that firms hold about the abilities of fraternity members, the fraternity's optimal admission policy only depends on μ , admitting a type (θ, μ) if and only if μ exceeds some critical cutoff. Hence, fraternity membership conveys no information to firms about θ . As a result, in equilibrium, both fraternity and non-fraternity members receive wage $E[\theta]$. Since wages do not depend on membership, it follows that only students with $n\mu \geq c$ apply. Figure 1.2 illustrates this uninformative signal case.

Although we do not explore it further, the case of completely uninformative signals about abilities highlights the gains that fraternities may achieve from an ability to commit to their admission policies. In particular, the fraternity would like to commit to excluding low ability students who have moderately high socializing values, and to accepting high ability students with lower socializing values. In practice, imperfect commitment devices that fraternities use include having university officials report the average GPA of members, and having the Greek council forbid fraternity participation to students with GPAs below some standard. This commitment induces a fraternity to weigh ability in admission, raising wages of members, and thereby raising the fraternity's payoff.

The central implication of Proposition 1 is that for fraternity membership to affect job market outcomes, firms must receive signals about student productivities that are noisy, but not perfectly so. Then, because more productive students tend to earn higher wages, fraternities value both productivity and fraternity socializing value, and will tradeoff between the two in admission. We now examine the choice problems of

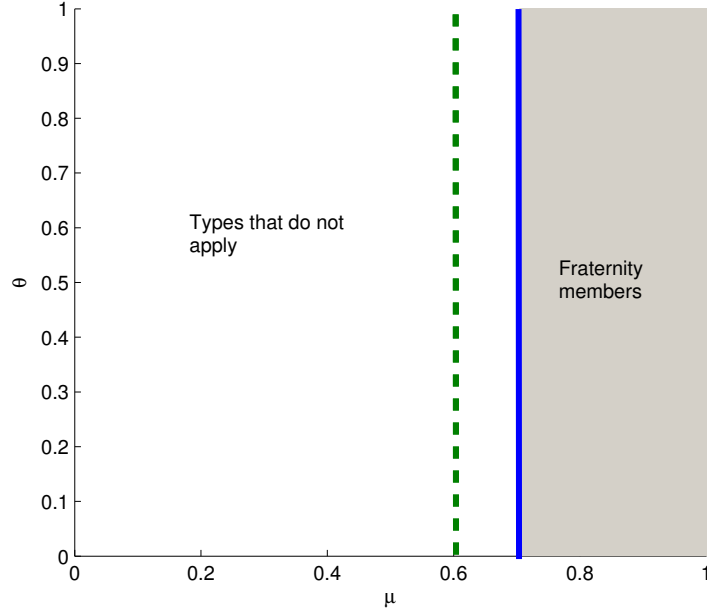


Figure 1.2: Equilibrium with Completely Uninformative Signals.

students and the fraternity in more detail.

1.2.1 Student's Problem

Students compare the expected payoffs from being a fraternity member and not, taking into account both the consequences for expected wages, and his fraternity socializing value. Optimization implies that a student type (θ, μ) applies for fraternity membership if and only if

$$E_{\tilde{\theta}} [w_C(\tilde{\theta})|\theta] + n\mu - c \geq E_{\tilde{\theta}} [w_{\bar{C}}(\tilde{\theta})|\theta]. \quad (1.2)$$

That is, a student applies to the fraternity either to obtain higher expected wages, or because his fraternity socializing value μ is sufficiently high.⁵ The following result follows straightforwardly.

Proposition 2 *If $a(\theta, \mu) = 1$, then $a(\theta, \mu') = 1$, for all $\mu' > \mu$.*

Proof. The expected wages of individuals (θ, μ) and (θ, μ') are the same, but a type (θ, μ') student gains strictly more utility from joining the membership. ■

⁵If (1.2) holds, then student type (θ, μ) applies even if he expects to be rejected by the fraternity. If students face positive costs of applying, then to reconcile the observation that some students apply, but are not admitted, one must integrate additional uncertainty/noise, so that a student does not always know whether he or she will be admitted. Such uncertainty complicates presentation and analysis, while providing limited benefits. Accordingly, we abstract away.

Corollary 1 *The equilibrium supply of fraternity applicants is summarized by a continuous function $\mu_A(\theta)$ such that a type (θ, μ) student applies if and only if $\mu \geq \mu_A(\theta)$.*

Proof. The result follows because expected wages of fraternity and non-fraternity members are continuous in θ . Therefore, student payoffs and hence choices are continuous in expected wages. ■

1.2.2 Fraternity's Problem

In any equilibrium, the fraternity offers membership to the set of students B that solves Problem 1. The sets A and B implicitly define the set of fraternity members $C = A \cap B$ and the set of nonmembers \bar{C} . Since the fraternity's payoff is increasing in the socializing values of its members, we have

Proposition 3 *For almost all (μ, θ) in $A \cap B$, almost all types (θ, μ') with $\mu' > \mu$ also belong to B .*

Proof. See the appendix. ■

We next show that we can extend this characterization to establish that the fraternity also wants to admit students who are more able as long as expected wages are increasing in θ ; and expected wages are increasing in θ if $w_C(\tilde{\theta})$ is increasing in $\tilde{\theta}$. As a preliminary step, we present an implication of the MLRP property on signals (see Paul R. Milgrom (1981)) for equilibrium wages.

Lemma 1 *Assume that $f(\theta|\tilde{\theta}) > 0$ on $[\underline{\theta}, \bar{\theta}]$, and that $f(\theta|x)$ satisfies the MLRP property. Fix a set $D \subset \Theta \times \tilde{\Theta}$ with $P(\tilde{\theta} = k|\tilde{\theta} \in D) < 1$ for all signals k . Let $Q(\tilde{\theta}) = E(\theta|\tilde{\theta}, D)$ be a firm's estimate of ability given signal $\tilde{\theta}$ and set D . Then from the perspective of a student with ability θ , his expected wage $E_{\tilde{\theta}}[Q(\tilde{\theta})|\theta, D]$, increases with his ability θ .*

Proof. See the appendix. ■

Proposition 4 *Suppose that the signals that firms receive about student abilities have the MLRP property. Then for almost all (θ, μ) if $b(\theta, \mu) = 1$, we have $b(\theta', \mu) = 1$ for almost all $\theta' > \theta$.*

Proof. By Lemma 1, the expected wage $E_{\tilde{\theta}}[w_C(\tilde{\theta})|\theta]$ is an increasing function of θ . The logical construction of Proposition 3 then applies. ■

Propositions 3 and 4 pin down the attributes of the set B of student types that the fraternity would admit. For example, if every student whom the fraternity would want to admit applies, then B is defined by a negatively-sloped curve in (θ, μ) space, $\mu_B(\theta)$: The fraternity admits almost every student type above (Proposition 4) and to the right (Proposition 3) of this curve (see Figure 1.3), i.e., $B = \{(\theta, \mu) | \mu \geq \mu_B(\theta)\}$, and $C = A \cap B$. Both $\mu_A(\theta)$ and $\mu_B(\theta)$ are continuous in θ , reflecting the continuity of expected wages in θ .

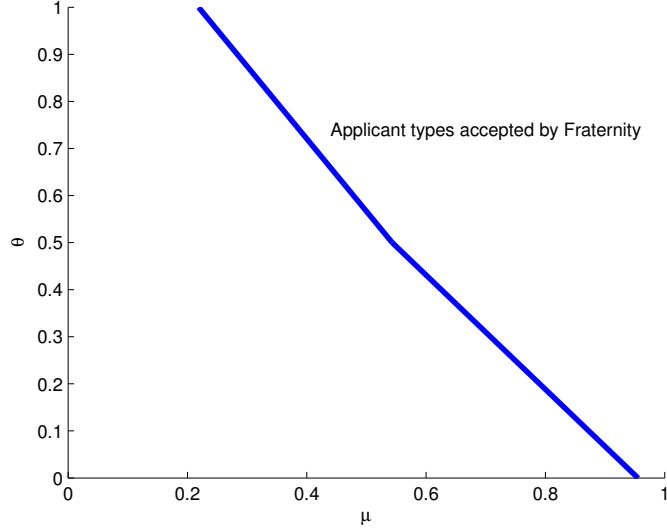


Figure 1.3: Unconstrained admissions set B .

More generally, for almost all θ where the fraternity's admission decision is not constrained by student application, i.e., for almost all θ with $\mu_B(\theta) > \mu_A(\theta)$, the fraternity trades off linearly between expected wage and fraternity socializing value in admission. That is, for θ_1, θ_2 with $\mu_B(\theta_j) > \mu_A(\theta_j)$, $j = 1, 2$, we have

$$W_1 E(w_C(\tilde{\theta})|\theta_1) + W_2 \mu(\theta_1) = W_1 E(w_C(\tilde{\theta})|\theta_2) + W_2 \mu(\theta_2).$$

That is, marginal contributions of these marginal types, $(\theta_1, \mu_B(\theta_1))$ and $(\theta_2, \mu_B(\theta_2))$ are equal.⁶

1.2.3 Existence of equilibrium

We first characterize when the “empty fraternity” is an equilibrium. In this “Groucho Marx” equilibrium, the fraternity would accept anyone who applies, but no one applies because firms believe that anyone who joins the fraternity has low ability $\underline{\theta}$ and hence would be given wage $w_C(\tilde{\theta}) = \underline{\theta}$. If no one joins the fraternity, then someone who generates signal $\tilde{\theta}$ receives wage $w_C(\tilde{\theta}) = E[\theta|\tilde{\theta}]$. Let $\underline{w} = E_{\tilde{\theta}}[w_C(\tilde{\theta})|\underline{\theta}]$ be the wage that a student with lowest ability $\underline{\theta}$ expects if he does not join the fraternity in this scenario.

Proposition 5 *Suppose that the signaling technology has a full support property, $f(\theta|x) > 0, \forall x$. Then an equilibrium exists with $A = C = \emptyset$ if and only if $n\bar{\mu} - c \leq \underline{w} - \underline{\theta}$.*

Proof. See the appendix. ■

⁶This result extends if we relax the structure on the fraternity's preferences, so that preferences over aggregate wages and socializing values are non-linear, $W(m(Ew_C(\tilde{\theta})), m(\mu|C))$. Then, the appropriate marginal derivatives, W_1, W_2 , evaluated at the aggregates, describe the indifference relationship for the fraternity.

If $n\bar{\mu} - c \leq \underline{w} - \underline{\theta}$, then pessimistic firm beliefs can support the empty fraternity equilibrium. However, if the inequality does not hold, then sufficiently inept students with high socializing values would prefer to join the fraternity because they also expect to receive low enough wages outside the fraternity that the maximum wage cost from joining the fraternity is more than offset by their high socializing values.

We next prove that an equilibrium always exists to this fraternity game, establishing a fixed point to a mapping from conjectured optimal student application and fraternity admission choices by firms to the best responses to those conjectures by students and fraternities. To do so, we exploit Propositions 2 and 4 and consider continuous student and fraternity choice functions $\mu_A(\cdot)$ and $\mu_B(\cdot)$, where a student type (θ, μ) is a member of the fraternity if and only if $\mu \geq \max\{\mu_A(\theta), \mu_B(\theta)\}$. Existence of equilibrium then follows from standard fixed point theorems.

Proposition 6 *An equilibrium exists to the fraternity game.*

Proof. See the appendix. ■

1.3 Three Signal Economy

To provide insight into the possible nature of equilibria to this fraternity game, we next consider an economy in which student productivities and fraternity socializing values are uniformly distributed on the unit square, i.e., (θ, μ) are uniformly distributed on $[0; 1] \times [0; 1]$, and students generate one of three possible signals, $\tilde{\theta} \in \{H, M, L\}$. This structure allows us to obtain explicit solutions for equilibrium outcomes. We will return to a more general signal and attribute distribution framework in the next section.

Here, we suppose that more able students with $\theta > 0.5$ generate either medium or high signals, where the probability of a high signal is linearly increasing in ability; and that less able students with $\theta < 0.5$ generate either low or medium signals:

$$\begin{aligned} \text{Prob}(H|\theta) &= 2\theta - 1, \theta > \frac{1}{2} \text{ and } 0 \text{ otherwise.} \\ \text{Prob}(L|\theta) &= 1 - 2\theta, \theta < \frac{1}{2} \text{ and } 0 \text{ otherwise.} \\ \text{Prob}(M|\theta) &= 1 - \text{Prob}(L|\theta) - \text{Prob}(H|\theta). \end{aligned}$$

This signal technology obviously satisfies the MLRP property. Its central feature is that a student with $\theta < 0.5$ hopes to get lucky and receive a medium signal, and thereby be indistinguishable from a student with $\theta > 0.5$ who unluckily receives a medium signal.

Let $w_C(\tilde{\theta})$ be the wage that a fraternity member who generates signal $\tilde{\theta}$ receives and $w_{\bar{C}}(\tilde{\theta})$ be the wage that a non-member who generates signal $\tilde{\theta}$. The expected wage of a student with ability θ who joins the fraternity is

$$E\left(w_C(\tilde{\theta})|\theta\right) = w_C(H) \text{Prob}(H|\theta) + w_C(M) \text{Prob}(M|\theta) + w_C(L) \text{Prob}(L|\theta).$$

An analogous expression describes wages of students who are not fraternity members.

The piecewise linear structure of the signaling technology implies that the expected wage functions are piecewise linear in θ with a single kink at $\theta = \frac{1}{2}$. It follows that the boundary describing the set of students that the fraternity would admit, where not limited by students' application decisions, is also linear with a kink at $\theta = \frac{1}{2}$. Since the difference in wages of fraternity members and non-members is linear with a kink at $\theta = \frac{1}{2}$, the boundary of the set of applicants to the fraternity, $\mu_A(\theta)$, is also linear with a kink at $\theta = \frac{1}{2}$. Therefore, the set of fraternity members, $\{(\mu, \theta) | \mu \geq \max\{\mu_A(\theta), \mu_B(\theta)\}\}$, is described by a continuous piecewise-linear function from $[0, 1]$ to $[0, 1]$ that has one or two kinks, where one kink is at $\theta = 0.5$, and the other (if it exists) is at the intersection of the fraternity and student cutoff rules.

One equilibrium is obviously the “empty fraternity”, but there are also more interesting equilibria. In particular, given Γ , we search for (i) an equilibrium in which the boundary $\mu_A(\theta)$ of A is everywhere to the left of the boundary $\mu_B(\theta)$ of B , i.e., where every student that the fraternity would want is admitted and (ii) an equilibrium in which $\mu_A(\theta)$ and $\mu_B(\theta)$ intersect, so the piecewise-linear function describing the frontier of the set of fraternity members has two kinks. This latter equilibrium is described by a system with thirteen unknowns (the slope and intercepts of the three lines plus the intersection point of the student and fraternity frontier, plus six wages) and thirteen equations (6 equations from the firm's problem — wages equal expected skill given signal realization and membership status, 4 equations from the fraternity and 3 equations from the students). We solve this system numerically for the associated equilibrium outcome, when it exists.

In our base parameterization, student utilities are $M + n\mu$, with $n = 0.18$, student time costs of participating in the fraternity are $c = 0.09$, the fraternity trades off between wages and socializing value according to $\frac{W_1}{W_2} = 1.1$, and the fraternity's capacity is $\Gamma = 0.35$. Figure 1.4 illustrates the unique “application-unconstrained” equilibrium. In this equilibrium firms have optimistic beliefs about the productivities of fraternity members, so that given any signal emitted by a student, his wage is higher if he is a member of a fraternity than if he is a non-member. As a result, in this equilibrium every student whom the fraternity would like to admit chooses to apply—and, indeed, because $w_C(H) = 0.8480 > w_{\bar{C}}(H) = 0.8143$, only very productive people with especially low socializing values choose not to apply (and while less productive

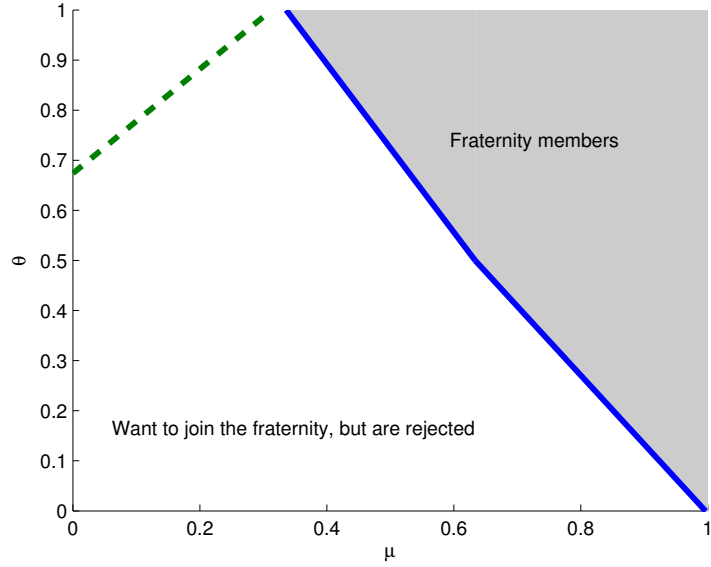


Figure 1.4: Application-unconstrained equilibrium.

students apply, most are rejected).

However, for exactly the same parameterization, there is also an “application constrained” equilibrium in which some students whom the fraternity would like to admit do not apply. Figure 1.5 depicts this equilibrium: the solid line denotes the fraternity’s cutoff rule, and dashed line denotes locus of students who are indifferent between joining the fraternity and not. In this equilibrium, firms hold more pessimistic beliefs about the abilities of fraternity members, so that higher ability students are more reluctant to join the fraternity. Intermediate quality students remain eager to join, and the fraternity’s composition is radically shifted to reflect this population. Comparing the fraternity’s cutoff line in Figure 1.4 with that in Figure 1.5 reveals that the fraternity is less “picky” when its choice set is constrained by the reluctance of able students to apply. Because able types $\theta = 1$ expect lower wages inside the fraternity than out, $w_C(H) = 0.7940 < w_{\bar{C}}(H) = 0.8555$, the fraternity attracts only a small fraction of able students, and the bulk of its members have intermediate abilities.

Figure 1.6 presents the expected wage that a student with ability θ would receive as a fraternity member and non-member for these two equilibria. Notice the crossing of wages in the application-constrained equilibrium. This reflects that while all lower ability student receive higher wages as fraternity members, higher θ students in the application-constrained equilibrium accept a direct loss in wage by joining the fraternity, for which they are compensated by high socializing values.

Note that in the application-unconstrained equilibrium, were we to increase Γ slightly, then all existing

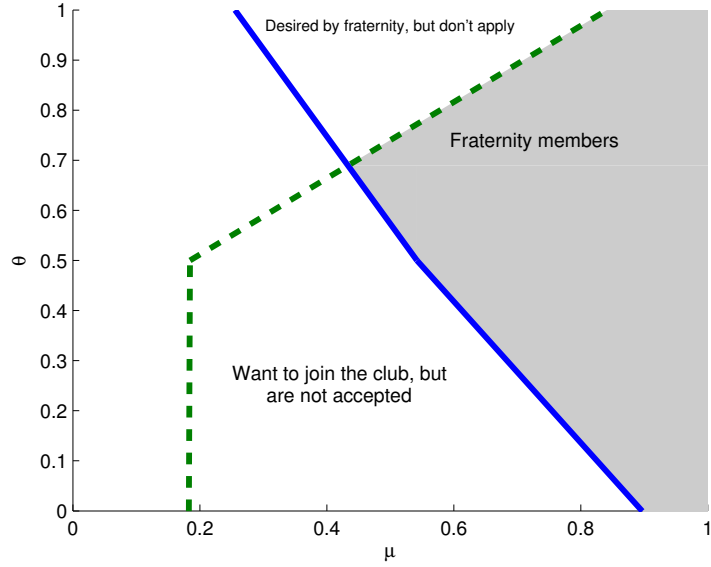


Figure 1.5: Application constrained fraternity game equilibrium.

members of the fraternity would still apply, and the fraternity’s payoff would be increased. This observation implies that were we to replace the fraternity’s capacity constraint with a strictly convex cost function of admitting more members, the fraternity would admit more members when beliefs of firms about member abilities are optimistic, thereby encouraging able students to apply.

Figure 1.7 reveals how the fraternity’s capacity affects equilibrium outcomes. Interestingly, raising capacity can *raise* the wages of able fraternity members. Essentially, when Γ is increased, the mix of students that the fraternity admits shifts slightly toward more able students with lower socializing values, i.e., toward students with higher θ s and lower values of μ . But this raises the expected wages of able students who join the fraternity. But then, able students are more willing to join—there is a significant increase in the measure of able students who apply to the fraternity. Notice also that as Γ increases, the slope of the boundary characterizing the application decision of less able students with $\theta \in [0, \frac{1}{2}]$ changes. This result reflects a change in the relative slope of wage functions: when, among students with $\theta \in [0, \frac{1}{2}]$, most of those with relatively high productivities are in the fraternity, then receiving the signal M and being outside the fraternity has a smaller premium than being in the fraternity and getting signal M (relative to receiving signal L in both cases).

1.4 Single-peaked Equilibria

Our three-signal setting shows that one possible equilibrium fraternity composition (see e.g., Figure 1.5) is where membership is a single-peaked function of student ability. Empirically (see Figure 1.8), we will

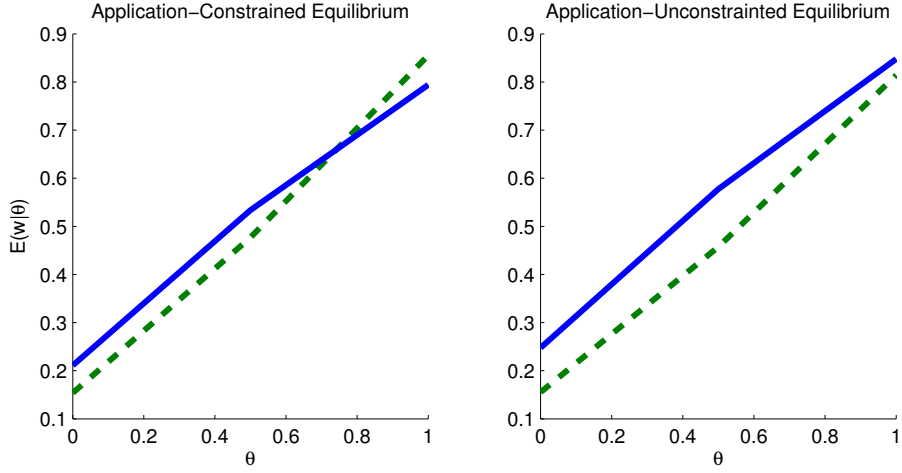


Figure 1.6: Equilibrium Wages. The solid line is the expected wage of a fraternity member (as a function of θ), and the dashed line is the expected wage of a nonmember.

see that conditioning on student ability, among sufficiently able students, the percentage who are members is a declining function of ability. This makes it important to understand when and how the single-peaked equilibrium emerges in a more general setting, and, in particular, to ensure that it is not the three-signal setting that underlies the single-peaked characteristic. Accordingly, we now identify sufficient conditions for membership to be a single-peaked function of student ability in *every* non-trivial equilibrium.

We say that an equilibrium set of fraternity members is *single-peaked* if:

1. $A \cap B \neq B$ and $A \cap B \neq \emptyset$.
2. The students' acceptance threshold $\mu_A(\theta)$ is increasing in θ for $(\theta, \mu_A(\theta)) \in B$.

The first condition says that there are students who would be accepted by the fraternity, $(\theta, \mu) \in B$, but choose not to apply. When the fraternity only values the future wages of its members, the first condition always holds as long as the costs c of joining are neither too small (possibly negative) that even the students with the lowest socializing values want to join, nor so large that even the students with the highest socializing values do not want to join. When the fraternity also values the socializing values of its members by enough that it does not accept the most able student with the least socializing value, then the necessary lower bound on c is higher.

We begin by showing that if the wage premium from membership in the fraternity falls with higher signals (recall, for example, Figure 1.6), then more able students are more reluctant to join the fraternity.

Lemma 2 *If $w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta})$ is a decreasing function of $\tilde{\theta}$ then the students' acceptance threshold $\mu_A(\theta)$ is an increasing function of θ .*

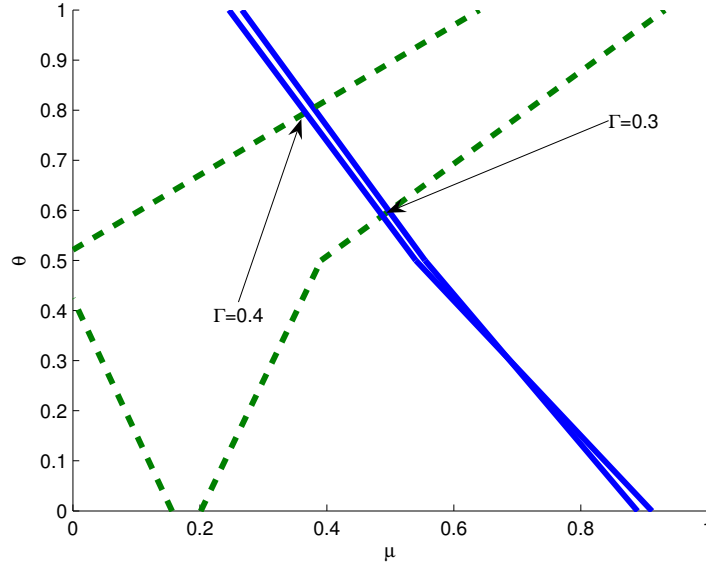


Figure 1.7: Application-constrained equilibrium for different capacities; arrows denote the intersection of fraternity's and students' choices in corresponding equilibrium.

Proof. See the Appendix. ■

We next establish conditions under which the single-peaked equilibrium emerges even when the fraternity only cares about the wages of its members, and not their socializing values. Then, it follows that the fraternity admits every students whose ability exceeds some cutoff, θ_0 , and further that as capacity Γ falls, θ_0 rises. We make two weak assumptions. The first technical assumption says that the support for signals is non-degenerate, and the second simply says that the cost of membership is such that students with the highest socializing value want to join, but those with the lowest socializing value do not.

Assumption 1 *Either the support for signals $\tilde{\theta}$ is finite, or the support of $f_{\tilde{\theta}}(\tilde{\theta}|\bar{\theta})$ is non-trivial.*

Assumption 2 *The cost c of joining the fraternity satisfies*

$$n\bar{\mu} + \bar{\theta} - E[\theta] < c < n\bar{\mu} + \bar{\theta} - E[\theta].$$

Proposition 7 *Suppose that Assumptions 1 and 2 hold and the fraternity does not care about the socializing values of its members. Then when the fraternity is small enough, the equilibrium is single-peaked. That is, there exists a $\underline{\Gamma} > 0$, such that for all $\Gamma < \underline{\Gamma}$, in every non-trivial equilibrium, fraternity membership is a single-peaked function of student ability.*

Proof. See the appendix. ■

The role of $\bar{\Gamma}$ is to provide a gross sufficient condition for a declining wage premium for fraternity membership—it ensures that the fraternity is sufficiently picky. The lowest ability fraternity members always receive a large wage premium, as their membership ensures that they are separated away from all lower ability types, and are mixed in with relatively higher proportions of more able types. The smaller is Γ , the higher is the wage premium for a given low ability type, due to the increased filtering out of lower ability students by the fraternity. High ability types gain less, because they only benefit from separation from low ability non-members, whom they may be unlikely to be confused with (as they are unlikely to generate the signals sent by low ability students), and fraternity membership lumps them in with intermediate ability students. When costs of membership are of an appropriate magnitude, they cause higher ability students to become increasingly reluctant to join, giving rise to the single-peaked equilibrium. That is, the fraternity’s filtering eliminates higher proportions of lower ability students, while on the high ability end, increasing proportions of higher and higher ability students choose not to join to avoid the increasing wage “penalty” for membership.

The logic of the proof of Proposition 7 extends immediately to settings where the fraternity cares about socializing values, so that $\mu_B(\theta)$ is a decreasing function of θ , but not so much that $\mu_B(\bar{\theta}) < \underline{\mu}$. As the weight W_2 that the fraternity places on socializing values increases, the analysis follows directly, albeit inelegantly, if we replace Assumption 2 with an assumption that we write implicitly in terms of equilibrium values:

Assumption 3 *Suppose that the cost c of joining the fraternity satisfies*

$$n\mu_B(\bar{\theta}) + \bar{\theta} - E[\theta] < c < n\bar{\mu} + \bar{\theta} - E[\theta].$$

In essence, if the fraternity places a sufficient weight on socializing values, it will not admit high ability students who have low socializing values, so it may already be filtering out the set of high ability students who are reluctant to join. As a result, a higher cost of membership may be required to support a single-peaked equilibrium.

Proposition 8 *Suppose that the fraternity places a positive weight on both socializing values and wages ($W_1, W_2 > 0$). Then when Assumptions 1 and 3 hold and the fraternity is small enough, the equilibrium is single-peaked. That is, there exists a $\underline{\Gamma} > 0$, such that for all $\Gamma < \underline{\Gamma}$, fraternity membership is a single-peaked function of student ability in any non-trivial equilibrium.*

Proof. The argument follows directly along the lines of the proof of Proposition 7, with the added structure on c guaranteeing that $\bar{\mu} > \mu_A(\bar{\theta}) > \mu_B(\bar{\theta})$. ■

Summing up, the single-peaked nature of the equilibrium is generated not by any particular specification of the signal structure that must be imposed to solve explicitly for equilibrium. Rather, the single-peakedness

derives only from the monotone likelihood property of the distribution of signals that firms receive about students combined with the conflicting interests of students and the fraternity that necessarily emerges whenever the fraternity is sufficiently selective and membership is costly.

- The monotone likelihood ratio property means that when the fraternity filters out low ability students, it is lower ability student types who gain more from fraternity membership, as most people who would generate low signals are rejected by the fraternity.
- The filtering out of low ability students by the fraternity, combined with the fraternity's trade off between future earnings and socializing values, initially leads to student participation in the fraternity being an increasing function of ability (for low θ types).
- The MLRP signal structure implies that higher and higher ability students gain less and less, or are even hurt in terms of wages by fraternity membership due to mixing in with less able students; and a cost-benefit calculation eventually causes higher ability students to become increasingly reluctant to join. This implies that when Γ is small, student participation in the fraternity eventually declines in ability.

1.5 Empirical Analysis

To see whether our model can reconcile the actual application and selection process of fraternities, we obtained data on the cumulative GPAs of the 8634 seniors at the University of Illinois in the fall semester of 2007 (excluding international students on temporary visas), and a random sampling of 701 seniors who were fraternity or sorority members. GPAs only reflect courses taken at the University of Illinois (i.e., omitting transferred courses), but the senior classification is based on all hours accumulated prior to the end of the fall, 2007 term.

Figure 1.8 presents the conditional probability that a student is a member of a fraternity given his or her GPA. This figure reveals that the conditional probability that a student with a low GPA of 2.0 is a fraternity member is less than 0.05, but that this probability more than triples for intermediate GPAs between 3 and 3.4, before falling by more than a third for students with high GPAs. Provided that the true distribution of ability is some monotone function of the GPA distribution, this relationship alone indicates that the data are inconsistent with pure signaling or pure screening models of fraternity formation. In particular, if fraternity membership has the properties of a signal, so that the benefit-cost tradeoff tends to be higher for more able students, then the fraction of students who are fraternity members would be *monotonically increasing* in GPA; and a similar relationship would emerge if fraternity membership is determined solely by the filtering out by fraternities of students with low abilities and socializing values. Moreover, the monotonic prediction

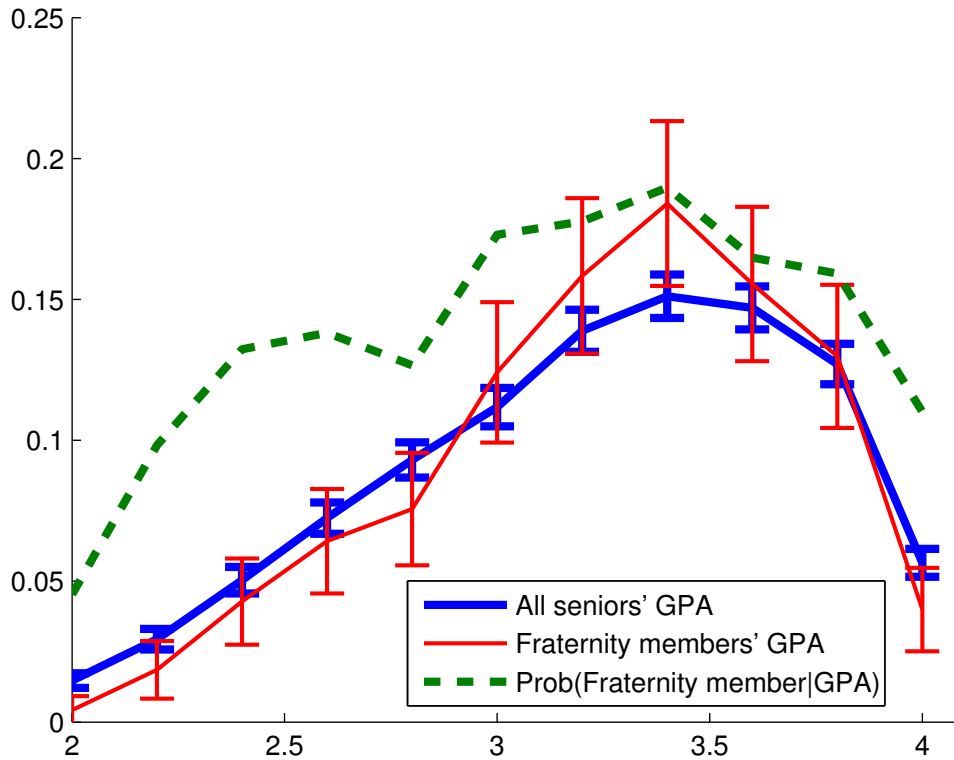


Figure 1.8: Conditional Distributions of GPAs

NOTES: The dashed line is the conditional probability that a student is a fraternity member given his or her GPA (rounded to the nearest 0.2). The thin line graphs the distribution of GPAs for 701 seniors who are fraternity members (fall 2007), and the thick line graphs the probability distribution for all 8634 seniors at the University of Illinois (fall 2007). The bars indicate 2 standard deviation confidence intervals.

of these signaling and screening theories would be *reinforced* if socializing values and ability were positively correlated in the population of students, rather than independently distributed as we have assumed. In addition, this single-peaked relationship is inconsistent with networking driving fraternity membership as long as ability and networking are either substitutes or complements in the wage determination process: if ability and networking are substitutes, then the fraction of students who are fraternity members should decline with GPA, as the expected networking benefit would be less for more able students; and if ability and networking are complements, then the fraction should increase with GPA.

In sharp contrast, this single-peaked pattern is precisely what emerges in the equilibrium to our fraternity game where able students are reluctant to join fraternities to avoid being tainted in labor market outcomes, while intermediate and less able students are eager to join, and the fraternity screens out most of the less able students, i.e., those who do not have high socializing values. This suggests that it is plausible to estimate our three-signal model of fraternity formation formally. That is, if we make the additional assumption that the

distribution of GPAs *corresponds* to the distribution of abilities in the population of students (rather than the distribution of abilities just being a monotone transformation of that of GPAs), then we can structurally estimate our model using this very limited dataset on grades and fraternity membership. Obviously, this assumption is a bit of a leap. However, it allows us to obtain estimates of primitive parameters that one can assess introspectively for plausibility, and, via these structural parameters, we can derive the estimated welfare impacts of the fraternity on students with different abilities and socializing values.

To be consistent with the premises of our three-signal model, we assume that θ corresponds to the quantile of the GPA distribution so that θ is distributed uniformly; and that fraternity-socializing values, μ , are independently and uniformly distributed.⁷

Step 1: Extracting observations of membership proportions from data. Fixing ability, our model indicates that fraternities admit higher μ types. This allows us to estimate $1 - \mu(\theta)$ directly: letting Φ be the event that a randomly-selected person is in the fraternity, the fraction of students with ability θ who are in the fraternity is

$$1 - \max(\mu_A(\theta), \mu_B(\theta)) = P(\Phi|\theta) = \Pr(\Phi) \frac{f_\theta(\theta|\Phi)}{f_\theta(\theta)},$$

where $\mu_A(\theta)$ and $\mu_B(\theta)$ are the cutoff rules of students and fraternity, $f_\theta(\cdot)$ is the density of θ , and $\Pr(\Phi)$ is the probability that a randomly-selected senior is a member of the fraternity. Our estimate of $\Pr(\Phi)$ is simply the number of senior fraternity members divided by the number of seniors, $\Pr(\Phi) = \frac{1345}{8634}$. To estimate $f_\theta(\cdot)$, we use the sample of all senior students with GPAs of at least 2,⁸ and we use the sample of fraternity members' GPAs exceeding 2 to estimate $f_\theta(\cdot|\Phi)$. We use the kernel estimates of the densities to smooth the conditional probability of being in the fraternity conditional on GPA (the dashed line in Figure 1.8). We then take 20 equally spaced θ s between 0.05 and 0.95 as our pseudosample, and evaluate $\hat{\mu}(\theta) = \max(\mu_A(\theta), \mu_B(\theta))$ from our smoothed conditional probability estimator at these points. This pseudo-sample is represented by the dots in Figure 1.9.

Step 2: Structural Estimation. In our three signal setting, given (a) the relative benefit of fraternity participation n , (b) membership cost c , (c) preference weighting of wages vs. μ by the fraternity, $\frac{W_1}{W_2}$, and (d) fraternity capacity Γ , there is a nonempty fraternity membership equilibrium that is described by a piecewise-linear function of θ that has three parts. Call these three linear segments the *lower*, *middle* and *upper* segments. Each segment is characterized by two parameters, the slope and the intercept. The lower and middle segments intersect at $\theta = 0.5$, and the middle and upper segments intersect at $\theta = k$ for some $k > 0.5$. The fraternity trades off linearly between expected wages and socializing values μ in admission,

⁷Relaxing these functional form assumptions, in the absence of other restrictions, can only allow us to better match the data.

⁸We drop the few students with GPAs below 2, as they are subject to screening by the University (and, indeed, only students with GPAs of at least 2 can graduate).

in particular for students with abilities above and below $\theta = 0.5$. Due to this fraternity optimization, the slopes of the lower and middle segments are linearly dependent with a coefficient that depends on wages, i.e., $\frac{W_1}{W_2} = \frac{b_1}{2(w_C(M) - w_C(L))} = \frac{b_2}{2(w_C(H) - w_C(M))}$, where b_1 is the slope of the club's cutoff rule below $\theta = 0.5$, and b_2 is the slope for $\theta > 0.5$. This means that four parameters describe the frontier of the fraternity, and there are four primitive parameters that determine this frontier in equilibrium. In practice, we estimate the frontier, and then back out the primitive parameters.

Because the fraternity optimization restriction is a non-linear function of wages, and wages are non-linear functions of the segment slopes and intercepts, we do not substitute this relationship into our objective function directly, but rather penalize deviations from this restriction, minimizing the residual sum of squares plus a quadratic measure of the distance from this restriction, where the penalty function is

$$10 [b_1(w_C(H) - w_C(M)) - b_2(w_C(M) - w_C(L))]^2.$$

That is, we estimate the three line segment slopes, k , and the intercept of the lower segment, imposing a sharp penalty for violations of the fraternity optimization restriction. That is, we have 5 parameters that define the fraternity frontier, 4 parameters to estimate and the fraternity optimization restriction. In our estimation:

- For a fixed k , we solve for the frontier that minimizes lasso-adjusted minimal sum of squares of deviations of the frontier from the data;
- Choose k to minimize the sum of squares;
- Verify that the penalty associated with the fraternity optimization restriction is small. In fact, our estimates yield a penalty value that is less than 10^{-10} , indicating that the estimated model is *very* close to being an equilibrium model.

Given any fraternity frontier, we can integrate to solve for the equilibrium wages for members and non-members, $w_C(L), w_C(M), w_C(H), w_{\bar{C}}(L), w_{\bar{C}}(M)$ and $w_{\bar{C}}(H)$. Integrating over the set of students inside the fraternity frontier pins down Γ . The upper line segment is determined by student indifference to membership. In particular, we can calculate the indifference conditions for students with abilities $\theta = 0.5$ and $\theta = 1$, who receive medium and high signals with certainty: $w_C(H) + n(a_3 + b_3) - c = w_{\bar{C}}(H)$ and $w_C(M) + n(a_3 + b_3/2) - c = w_{\bar{C}}(M)$, where a_3 is the intercept of upper segment and b_3 is the slope. Solving these linear equations yields n and c . Finally, we recover $\frac{W_1}{W_2}$ from lower line segment, exploiting $\frac{W_1}{W_2} = \frac{b_1}{2(w_C(M) - w_C(L))}$.

Step 3: Non-structural Estimates. We contrast estimates from this structural model with those from a non-structural model that identifies the slopes and intercepts of the three line segments and second kink

Parameter	OLS		Structural	
	Estimate	95% Conf. Int.	Estimate	95% Conf. Int.
Constant	0.8815	(0.8564, 0.9135)	0.8825	(0.8505, 0.9067)
Slope bot. line	0.1383	(0.0521, 0.2459)	0.1435	(0.0376, 0.2131)
Slope mid. line	-0.1020	(-5.1359, 0.6049)	0.1422	(0.0353, 0.2072)
Slope top line	-0.2342	(-0.3973, -0.0243)	-0.1201	(-0.3178, -0.0640)
Intersection point	0.8461	(0.5023, 0.8673)	0.5000	(0.5018, 0.7754)
n			0.2771	(0.1193, 0.5312)
c			0.2281	(0.0895, 0.4449)
c/n			0.8234	(0.7141, 0.8147)
W_1/W_2			0.2227	(0.0565, 0.3346)
Γ			0.1563	(0.1546, 0.1577)
Discrepancy from equil.	0.1410	(0.0012, 65.7960)	8×10^{-11}	$(3 \times 10^{-11}, 2 \times 10^{-05})$

Table 1.1: Estimates

that best describe the fraternity membership data (i.e., our non-structural estimates minimize the sum of squared errors). We obtain our non-structural estimates using a two-step estimation procedure in Matlab. For a given kink location, the procedure finds the slopes of the cut-off rules that minimize the sum of squared errors. The procedure then identifies the kink location that minimizes the SSE.

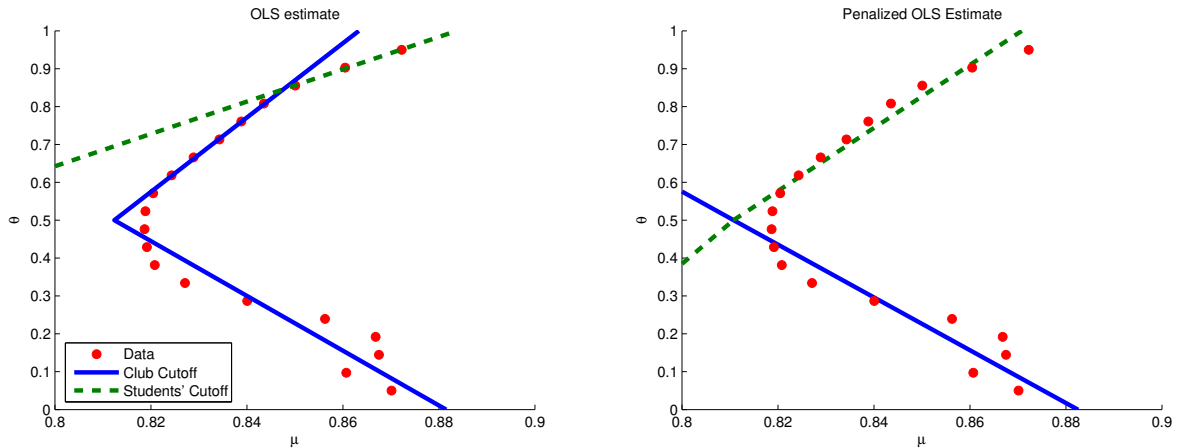


Figure 1.9: Estimated Unconstrained Model and Structural Model Fits

The first panel of Figure 1.9 presents the estimated cut-off rules of the unconstrained model, where we do not impose consistency of wages with the implied estimates of the ability distributions of members and non-members. This fit is *far* from an equilibrium: most obviously, the fraternity’s “cut-off rule” is not a monotonically decreasing function of θ . The panel on the right presents our structural estimates of the frontier describing fraternity membership. An F-test⁹ indicates that the differences between the structural and non-structural fits are only marginally statistically significant. However, the estimate of the intersection

⁹The F-value is 2.93, (1,13) degrees of freedom, p-value of 0.11; this statistic, however, presumes the normality of errors, but given the nonlinearity of both the model and the equilibrium restrictions, this presumption is likely to be violated. The asymptotic distribution of this test, which does not depend on normality of errors, is $\chi^2(1)$, with a p-value of 0.08.

point k is at the boundary, 0.5. suggests that the particular form of the three signal structure is misspecified. The confidence interval from bootstrapped samples however suggests that the signal structure is not violently inconsistent with the data.

Our structural estimates can be used to back out estimates of the primitives for students: the cost c is 0.23, or about 46% of the unconditional expected wage, and the fraternity socializing parameter is $n = 0.28$. We bootstrap the estimator to obtain 95% confidence intervals. While the confidence intervals for n and c are wide, the fraction $\frac{c}{n} = 0.82$ has a tight 95% confidence interval of $[0.71, 0.82]$; and it is this ratio that determines whether a student gains a net utility benefit from joining the fraternity in a full information setting where firms know a student's ability. Figure 1.9 shows that most fraternity members are above this threshold, i.e., most fraternity members have socializing values of $\mu > \frac{c}{n} \approx 0.82$. Thus, relative to a full information setting, the wage-setting mechanism impedes the efficiency of fraternity participation: there are too many fraternity members with intermediate abilities and too few low and high ability students with high socializing values. Our estimate of the fraternity's relative weighting on member wages versus socializing values, $\frac{W_1}{W_2}$ of 0.22 has a wide confidence interval, but the fraternity's capacity is precisely estimated.

Our estimates allow us to explore how the presence of the fraternity affects the welfare of different student types. Figure 1.10 presents welfare gains and losses of different student types relative to a setting in which there is no fraternity [or equivalently relative to the "empty" fraternity equilibrium]. The solid line divides the population of student types into those who benefit and those who are hurt, and the darker is the shade in the figure, the more the fraternity's presence hurts/benefits less a student. The figure reveals that all fraternity members actually are made better off by the presence of the fraternity, gaining from the socializing values of fraternity membership. In addition, able types, $\theta > 0.61$ who are not members gain because they receive higher market wages—firms believe that most highly productive types do not join the fraternity.

Figure 1.11 contrasts student welfare in the equilibrium with that which obtains when firms ignore the information in fraternity membership, or equivalently where the fraternity does not have better information about θ than firms, and hence only weighs socializing values in admission. Two groups of student types benefit when the fraternity weighs both expected wage/ability and socializing values in admission: (i) low ability types with high enough socializing values that they are admitted benefit from wage gain associated with being mixed in with more able types; and (ii) high ability types with lower socializing values who do not join benefit from the higher wages due to the partial separation from mediocre ability types generated by the fraternity admission process. Low ability types with moderately high socializing values are hurt the most, as they both would gain socializing values were the fraternity not to weigh ability, and they are punished by lower wages due to their exclusion. The other group hurt consists of high ability/high socializing value

Parameter	Fraternities		Sororities	
	Estimate	95% Conf. Int.	Estimate	95% Conf. Int.
Constant	0.8597	(0.8461, 0.9028)	0.8842	(0.8436, 0.9316)
Slope bot. line	0.0360	(0.0070, 0.1642)	0.1700	(0.0660, 0.3204)
Slope mid. line	0.0337	(0.0068, 0.1572)	0.1593	(0.0594, 0.3061)
Slope top line	-0.1397	(-0.4324, -0.0186)	-0.2152	(-0.3601, -0.0774)
Intersection point	0.6886	(0.5021, 0.8447)	0.5632	(0.5041, 0.5878)
n	0.2109	(0.0529, 1.3353)	0.2754	(0.1134, 0.5969)
c	0.1725	(0.0330, 1.1294)	0.2204	(0.0798, 0.4942)
c/n	0.8179	(0.6367, 0.8485)	0.8003	(0.6836, 0.8295)
W_1/W_2	0.0543	(0.0106, 0.2525)	0.2636	(0.1009, 0.5148)
Γ	0.1496	(0.1474, 0.1511)	0.1637	(0.1608, 0.1662)

Table 1.2: Estimates for Gender-Specific Models

students who join in both environments, but are tarred by association with lesser types when the fraternity weighs ability in admission, and therefore receive lower wages.

Similar picture can be seen if we control for gender — see Table 1.2. The results seem different from the gender-blind version, but both argue against a simple signaling model. It seems that selection matters more in fraternities (k is higher), the monetary cost of participation is lower in fraternities, but the payoff for social skills is bigger in sororities. This can be extended further (controlling for age, race, income and so on), but not much more new information will be obtained without more data.

An alternative estimation technique, proposed by George Deltas, delivers similar results. Instead of minimizing the sum of squares of deviations of data-based pseudopoints from the frontier implied by structural parameters, we maximize the likelihood of observations:

- We assume we can use each student's quantile of GPA in the whole students' sample as a proxy for θ .
- We split the θ axis into intervals: we use 69 equispaced intervals between 0.05 and 0.99, $[0, 0.05]$ and $[0.99, 1]$; use i to index these intervals.
- We estimate the probability that a student becomes a member of the fraternity \tilde{p}_i and a probability that a student does not become a member of the fraternity \tilde{q}_i from our sample.
- We parameterize the shape of the fraternity frontier, and solve for implied probabilities p_i and q_i of a student from i th interval of being a fraternity member and not being a fraternity member, respectively.
- By changing parameters of the shape of the fraternity frontier (and therefore p_i and q_i), we minimize

$$\sum_i -(\tilde{p}_i \ln p_i + \tilde{q}_i \ln q_i).$$

- We penalize the likelihood function for deviating from the equilibrium frontier, very much like in the previous method.

The results can be seen in Table 1.5, they are not too much different from the results of the OLS-based method. Identification of the intersection point of two cutoff rules requires a very fine grid of intervals for the second method, so that a method using an acceptably fine dissection is significantly slower than the original method. On the other hand, this weak dependence on the exact position of the intersection helps to get more stable estimates: one can see that now all estimates are inside the confidence intervals, although 0.5, the θ point where relevant signals change, still binds for a lot of bootstrap pseudosamples.

We conclude our empirical analysis by re-emphasizing caveats and limitations of this analysis. First, as we highlighted earlier, while the estimated model and welfare impacts seem plausible, the GPA distributions may differ from those for abilities, and our estimates hinge on their equivalence. Indeed, one could contemplate the possibility that fraternity membership in and of itself alters the distribution of GPAs. For example, it could be that not only does ability influence GPAs, but so does fraternity membership—fraternity cheat sheets may help low ability students, while a fraternity party environment may make it difficult for high ability students to study. Ideally, one would obtain measures of ability such as high school grade or ACT scores that are not affected by fraternity membership. We also note that to show that the ability distribution in the fraternity that we obtain theoretically is not driven by direct factors, we assume that firms do not value fraternity socializing values and that socializing values are uncorrelated with ability. Still, our empirical finding that high GPA students are reluctant to join fraternities indicates that, in practice, firms cannot value those skills by too much, and that the correlation between ability and socializing value cannot be too high.

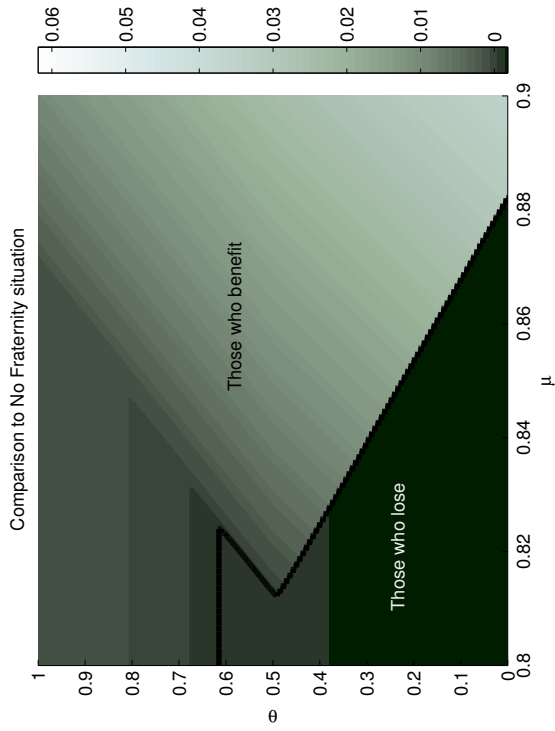


Figure 1.10: Impact of Fraternity on Student Welfare

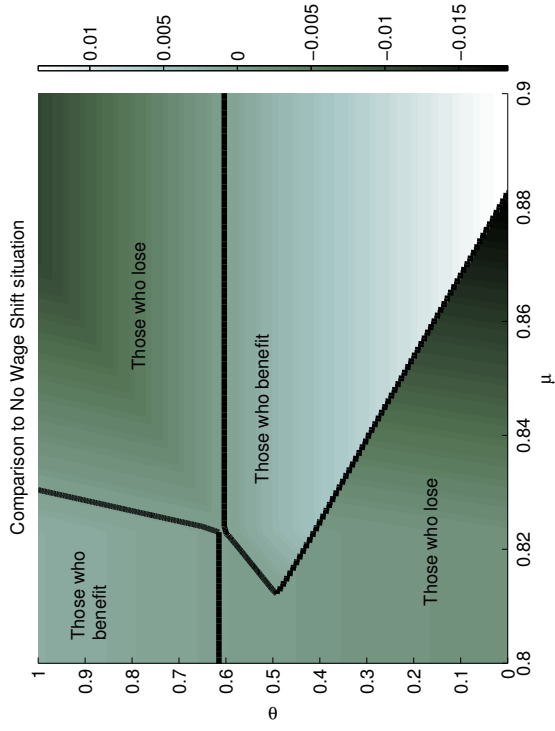


Figure 1.11: Impact of Wage Redistribution on Student Welfare

∞

Parameter	Together		Fraternities		Sororities	
	Estimate	95% Conf. Int.	Estimate	95% Conf. Int.	Estimate	95% Conf. Int.
Constant	0.8882	(0.8618, 0.9135)	0.8687	(0.8495, 0.9012)	0.9067	(0.8661, 0.9440)
Slope bot. line	0.1565	(0.0574, 0.2338)	0.0632	(0.0053, 0.1673)	0.2489	(0.1194, 0.3720)
Slope mid. line	0.1540	(0.0548, 0.2276)	0.0617	(0.0053, 0.1628)	0.2450	(0.1176, 0.3652)
Slope top line	-0.1314	(-0.2620, -0.0517)	-0.1051	(-0.3587, -0.0272)	-0.1902	(-0.3253, -0.0730)
Intersection point	0.5084	(0.5000, 0.6937)	0.5866	(0.5000, 0.8394)	0.5000	(0.5000, 0.5253)
n	0.2985	(0.1568, 0.5922)	0.2646	(0.0653, 0.8033)	0.2995	(0.1594, 0.6459)
c	0.2462	(0.1219, 0.4984)	0.2200	(0.0448, 0.6805)	0.2426	(0.1234, 0.5378)
c/n	0.8248	(0.7635, 0.8410)	0.8314	(0.6546, 0.8481)	0.8100	(0.7341, 0.8305)
W_1/W_2	0.2431	(0.0860, 0.3706)	0.0961	(0.0082, 0.2581)	0.3958	(0.1835, 0.6135)
Γ	0.1553	(0.1543, 0.1559)	0.1485	(0.1474, 0.1494)	0.1628	(0.1610, 0.1640)

Table 1.3: Estimates for Gender-Specific Models: Likelihood-Based

1.6 Conclusion

On first impression, it is not clear why fraternity or sorority membership should matter for labor market outcomes—fraternity activities seem to have little to do with skill development for future careers. Nonetheless, resumes regularly highlight fraternity membership, suggesting that membership augments the other signals that employers use to evaluate a person’s productivity. Our paper provides insights into when fraternity membership matters for labor market outcomes. We first show that if firms can either evaluate student productivities perfectly, or are completely incapable of screening job applicants, then fraternity membership has no impact on labor market outcomes. Otherwise, fraternity membership matters. In particular, we identify two equilibria in which fraternity membership is valued by some students for labor market outcomes. In one equilibrium, optimistic beliefs by firms about the abilities of fraternity members lead to higher wages for fraternity members than non-members. As a result, everyone whom the fraternity would like to admit chooses to pledge. We also identify an equilibrium in which able students are harmed in the labor market by fraternity membership, but less able students benefit. In this equilibrium, most fraternity members have intermediate abilities: less able students apply, hoping to be mixed in with better students, but are rejected unless they have high fraternity socializing values, while very able students who lack high socializing values do not apply to avoid being tainted in labor market outcomes due to being mixed in with less able fraternity members. We find that this latter equilibrium can reconcile the qualitative features of the ability distributions of fraternity members and non-members at the University of Illinois.

While we pose our analysis in the context of fraternities, the central economic story extends with some variations to filtering by other organizations. For example, ROTC (reserve officer training corps) may value both intellectual ability and leadership skills that firms value, but also physical fitness that does not contribute productively in many occupations. As a result, even were ROTC not to directly build skills of its officers, our model indicates that firms may rationally weigh ROTC membership positively in their evaluations of job-seekers.

Chapter 2

Equilibrium Sovereign Default with Endogenous Exchange Rate Depreciation

2.1 Introduction

Current literature on endogenous defaults makes a point that the level of debt burden a country can sustain is limited by the country's income. Starting with Arellano (2008), endogenous default is modeled as a choice of a sovereign between repaying the interest and staying in good relations with abroad and defaulting and suffering some sort of a penalty. Bulow and Rogoff (1989) showed that purely reputational penalty in a form of temporary exclusion from credit markets is not enough to sustain a positive debt level in equilibrium. In the world of Arellano (2008) positive debt level is sustained by explicit progressive penalty to GDP of the defaulting country. Necessity for penalty progressiveness is levied away by Aguiar and Gopinath (2006), which assumes some persistence of income shocks. Yue (2010) considers Nash bargaining in settling the default episode, and Bi (2008) makes it a repeated bargaining game to explain the uneven duration and the penalty size. The income penalty channel is motivated by an empirical finding that output falls after default on average. Tomz and Wright (2007), however, finds the link between "bad times" and default to be "surprisingly weak" — about 40% of the default episodes end up with above the trend output.

Current theoretical literature tends to model interaction with foreign countries for imperfect insurance against future recessions, and treats defaults in isolation from financial crises. However, a lot of empirical studies demonstrate that international trade is severely affected by the default. Rose (2005) documents that default reduces international trade by 8% for an extended period after default. He speculates that trade suffers after default because international flows require short-term financing: in the aftermath of default, trade partners will not extend this credit and raise the cost of cross-border trade. This conjecture is consistent with Arteta and Hale (2008), who demonstrate that private firms find international credit scarce after a sovereign default. Aside from disruptions to financing, the diminishing trade may be due to already imperfect trade mechanisms. Exporters are at the whim of other country's trade officers, who can greatly increase the cost of trading goods by overzealous customs inspections or other non-tariff trade barriers. The time cost, deriving from uncooperative trade officials, can add 10-30% to the cost of imports, as estimated by Hummels (2001).

Many of these costly bureaucratic barriers must be fungible and can be made more stringent for less-favored trade partners. Furthermore, many developing economies depend on favorable trade agreements and sourcing partnerships with the developed world.¹ Without favored market access or the use of a trade partner transport infrastructure, the effect on the cost of exporting goods is similar to additional “iceberg” costs.

2.1.1 Patterns of Trade in Default

De Paoli and Hoggarth (2006) explore defaults since 1975 and find a strong link between currency crises and sovereign default. They propose an informal explanation for this observation related to nominal rigidities and potential central bank insolvency. Here we expand on this finding: Table 2.1 shows that even when depreciation might not qualify as a crisis, defaulters usually experience a decline in their terms of trade. While other literature often keys on “hot” flows of funds and large changes to the nominal exchange rate, the real depreciation actually accounts for most of the change in relative prices. Sovereign default also brings adjustments to real trade volumes.

Table 2.1 presents the gross changes over a year for defaulters’ nominal effective exchange rate (NEER), real effective exchange rate (REER), export price and the fraction of expenditure on imports. Notice that the median defaulter has nearly the same depreciation in real exchange rate as in nominal exchange rate after a year’s time. With higher frequency data, the NEER depreciation tends to be sharper, but the REER quickly catches up. In some cases, the nominal change is greater than the real, but in many of these instances, narrative evidence suggests mismanagement by the monetary authority or financial market overreaction. Table 2.1 reveals that nominal appreciation is rather uncommon.

Gross change	NEER	REER	Export Prices	Imports/GDP
Mean	0.8492	0.8900	0.9184	0.8968
Median	0.9102	0.9220	0.9075	0.9211
Pr <.2	0.6835	0.7861	0.8291	0.7773
Pr <.4	0.8729	0.8688	0.8984	0.8978
Pr <.6	0.9505	0.9429	0.9269	0.9436
Pr <.8	1.0168	0.9755	1.0068	1.0070

Note: Exchange rate listed as foreign goods per home currency unit. Our data comes from the IMF’s IFS database and looks at defaulters since 1975, as identified by Beers and Chambers (2003).

Table 2.1: One year effect of sovereign default.

Countries in default suffer a consistent drop of about 10% in the price of their exports in foreign markets; the distribution of price changes is more tightly clustered than the depreciation distribution. Because the REER weights their trading partners by volume of trade, it suggests that the fall in export prices is more pronounced with a country’s more active trading partners. Defaulters receive fewer imports in return for

¹195 active bilateral trade agreements have been reported to the World Trade Organization by January 2009, any of which could be terminated in response to a financial episode and would effectively increase tariffs faced by a country’s exporters.

the same number of export goods, a fall in export price, which might stimulate domestic households to substitute away from foreign goods, further affecting the exchange rate.

Even at high frequency, the depreciation is abrupt and follows the same trend as above, with REER slightly trailing NEER. In Figure 2.1, we show the exchange rate dynamics in two recent defaults, Ukraine and Paraguay. Note that in the Ukraine, real and nominal exchange rates move almost indistinguishably. This suggests that post-default currency depreciation is not only due to shaky sovereigns who impel hot capital to flee — instead something fundamental happens to cross-border exchange of real goods.

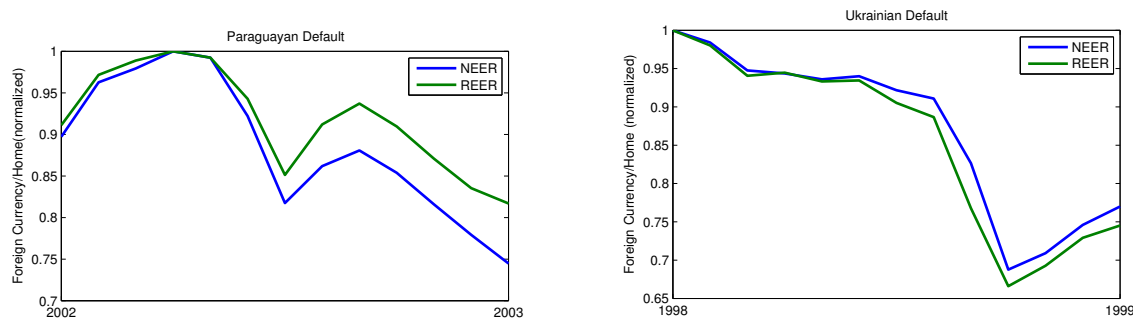


Figure 2.1: Dynamics of exchange rates in Paraguay and Ukraine, nominal and real. Exchange rates are normalized so that maximum value on sample is 1.

The data also reveal that households with bankrupt governments also spend relatively less on imports, which have become more expensive following the price changes. This fact is an amalgam of competing forces: the substitution effect would decrease imports, and the income effect would also diminish the demand for imports, with imports as normal goods. On the other hand, average income is lower in default. The falling share of imports suggests that this latter effect is not so strong. In our model too, imports fall as a share of spending. Though Tomz and Wright (2007) suggest that income is not a prerequisite of default, it does seem to reliably change trade flows.

2.1.2 Trade Channel Penalty

Bulow and Rogoff (1989) suggest that trade sanctions might sufficiently discourage default as to support realistic levels of debt. They justify such trade disruptions by speculating that a country might rely on its reputation to maintain trade flows, and it would be damaged by a default episode. They were not, however, modeling the trade channel penalty, or making quantitative statements about its implications. Investigating the role of trade penalties further, Wright (2001) discusses the similarities between penalties to trade flows, assets or income. In some environments, there is an equivalences in that any of these penalties can create a welfare loss that dominates the gain from a country defaulting.

This paper introduces a trade channel penalty, a second vehicle in which default creates costs for a sovereign. Post-default, we consistently observe real depreciation and substitution away from imports. A model without trade in goods cannot account for the composition effect of default decision on international trade. We argue that default is punished by a deterioration to the terms of trade. At its extreme, this would be an absolute embargo, which was not uncommon historically when countries refused to pay their creditors. In the 1861 Mexican default, creditors actually seized the port of Veracruz (see Todd (1991)). Circumnavigating gunboats certainly increased the “iceberg costs” on Mexican exports. However, Tomz (2007) argues that even in the 19th century heyday of gun boat diplomacy, outright military force only played a minor role in punishing defaulters.

The contribution of this paper is to develop a richer model in which there is trade in both borrowing markets and consumption goods. In particular, individuals care about consumption of both a home, domestic, good and a foreign, imported good. Our model can address the extent to which sovereign default leads to real exchange rate depreciations. The sovereign we model maximizes the citizens’ total future expected utility. Sovereign trades with abroad; equilibrium rate of exchange of foreign goods for home production is treated as an exchange rate. Sovereign borrows from abroad in hard times promising to repay in future. When sovereign chooses to default, foreign traders get upset, effectively imposing iceberg costs on imports, which makes imports more expensive. Benevolent sovereign, responding to worsening of terms of trade, reallocates the citizens’ consumption bundle. Thus, in equilibrium, default episode can affect the exchange rate.

Our contribution is to integrate trade in goods into models of sovereign default. Now default affects not only access to capital markets, but also terms of trade, and we can address how default affects exchanges rates, and how default affects the market for exportable goods. This gives rise to a richer default problem for the sovereign, as the sovereign now cares about not only how default affects the level of consumption of domestic consumers, but also how default affects the composition of consumption for domestic consumers. As a result of calibration to Argentina at the brink of 2000s, we obtain reasonable predictions of post-default depreciation and import substitution. We find that default decision is sensitive to preference parameters over imports, thus suggesting that more internationally integrated countries will tend to default less. Overall, we find that trade channel penalty explains a lot of default decision behavior in the income channel literature plus it explains international trade adjustment, unfeasible in previous models.

2.2 The Model

Our model of default extends the models of Eaton and Gersovitz (1981) and Arellano (2008), integrating a commodity space with both domestic and foreign goods. We describe a small open economy in which a government internalizes its citizens preferences over domestic goods, c_t , and imports, m_t . These goods are imperfect substitutes with constant elasticity, $\frac{1}{1-\kappa}$ in our formulation. The relative price of exports is the exchange rate, e_t . Imports are exchanged for exports, x_t by import firms according to $m_t = f(x_t)$. Their profit, Π_t , goes back to households.

Asset markets are incomplete with one period bonds serving as the only insurance against shocks to the income stream, y_t . The sovereign borrows b_t on behalf of its citizens from an international market paying a coupon q_t . Debt contracts are not enforceable, so the country may default. As punishment, the country suffers a deterioration in terms of trade and financial autarky for a random period of time.

The domestic country maximizes

$$U(c_t, m_t; \kappa, \alpha), \quad (2.1)$$

subject to the income process and budget constraint

$$c_t + e_t m_t + b_t = y_t + q_t b_{t+1} + \Pi_t, \quad (2.2)$$

$$\log y_t = \rho \log y_{t-1} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, s^2). \quad (2.3)$$

The importer transforms exports to imports according to

$$m_t = f(x_t) : f'(\cdot) \geq 0, f''(\cdot) \leq 0. \quad (2.4)$$

In particular, we specify the import export technology to be $f(x_t) = \theta_1(x_t - \theta_0)^\theta$. This technology is a reduced-form representation of the rest of the world's demand for domestic goods. For $\theta < 1$, the home country's exports receive diminishing returns in terms of imports. This is equivalent to assuming diminishing marginal utility of the country's goods to the foreign consumers. θ_0 allows for fixed costs to exporting, though in our calibration we find its value to be small.

Certainly, we could motivate the demand coming from the rest of the world (ROW) with a structural formulation for preferences. However, from the point of view of the small open economy, only the ROW demand can be observed, and this demand function is sufficient to pick the optimal quantity of exports.

Both estimating and imposing an explicit form for the ROW utility seems treacherous. From an economist’s perspective, calibrating the preferences of the ROW is daunting because one must aggregate structural preferences, but the weighting of the relative importance of a trade partner is only observed in equilibrium. Aesthetically, a full optimization by the ROW might be preferable, but we choose to keep our model focused and parsimonious; our contribution is not on why countries trade, but rather on why they default.

Much of our analysis would go through with a simpler linear technology, $f(x_t) = \theta_1 x_t$. With a linear technology, the trade channel penalty would cause depreciation in default, but the depreciation would be one-for-one with the size of the penalty. We would essentially be directly imposing the size of depreciation, instead of allowing it to be determined endogenously by preferences. Our preferred form better matches the stylized facts. In particular, (with a linear technology) the depreciation would be the same size every time, but Table 2.1 reveals that this is not the case.

2.2.1 Timing

Time is infinite and discrete. At the beginning of each period, a country’s state variables — income, borrowing to repay, and default status — are common knowledge. The first choice is whether to default; then country chooses consumption, trade and borrowing policies. If the country is in good standing, it faces a coupon price schedule from its creditors who will charge a premium over the fixed world interest rate based upon the probability of default in the next period.

If the country defaults, then it experiences a terms-of-trade deterioration before its trade policy is set. The equilibrium exchange rate clears the international trade markets. As in Eaton and Gersovitz (1981), countries in default are in financial autarky, and their bond position reverts to zero. Unlike the aforementioned article, countries can leave this defaulted state each period with an exogenous probability of being “forgiven”.

2.2.2 Representative Household and Government

The sovereign government internalizes the problem of the representative household. In particular, household preferences take the form $u(c_t, m_t; \kappa, \alpha) = [\alpha c_t^\kappa + (1-\alpha)m_t^\kappa]^{\frac{1-\sigma}{\kappa}} / (1-\sigma)$. We will formulate the representative household problem recursively. Given income realization, y , bond position, b , and the indicator of whether it is being punished, taking as exogenous the price of bonds, $q(\cdot, \cdot)$ and imports, $e(\cdot)$, the government solves two subproblems, one if it decides to default and one if it does not default. The solution to the non-default problem is summarized by the value function $V(b, y)$ and the defaulter’s problem has value function $W(y)$. The default decision is the argmax of these two functions, and the government’s value function, $U(b, y)$ is the envelope over the subproblem value functions, $V(\cdot, \cdot), W(\cdot)$. Formally, the government solves the **Household**

Problem:

$$U(b, y) = \max_{h \in \{0,1\}} hW(y) + (1-h)V(b, y),$$

where $h \in \{0, 1\}$ indicates default. Conditional on not defaulting this period, the **Household's Problem In Good Standing** is:

$$\begin{aligned} V(b, y) &= \max_{c, m, b'} u(c, m) + \beta EU(b', y'), & (2.5) \\ \text{s.t.} & \\ c + em + b &= y + q(y, b')b' + \Pi, \\ \ln y' &= \rho \ln y + \epsilon, \quad \epsilon \sim \mathcal{N}(0, s^2). \end{aligned}$$

The debt discount $q(y, b')$ adjusts the price of borrowing to accommodate the probability of default. By choosing $b' = b_0$, the country accepts the contract that gives it $q(y, b_0)b_0$ units of home good this period and takes away b_0 units of home goods next period. Using discounts instead of returns permits banks to decide not lend at all, which may happen when default is nearly certain because of an extremely high level of debt, which imply an infinite interest rate.

If country chooses to default, then its value function is an optimal solution to **The Household Problem In Default**:

$$\begin{aligned} W(y) &= \max_{c, m} u(c, m) + \beta E(\phi W(y') + (1-\phi)U(0, y')), & (2.6) \\ \text{s.t.} & \\ c + em &= y + \Pi, \\ \ln y' &= \rho \ln y + \epsilon, \quad \epsilon \sim \mathcal{N}(0, s^2). \end{aligned}$$

2.2.3 Importers

The importers face a trade demand given by equation (2.4) from the rest of the world. Note, that the country does *not* face a perfectly elastic demand from the rest of the world. Instead, this is a “small country model” only in that the decisions of the country do not affect the demand of the rest of the world. In case of default, equation (2.4) is shifted to generate the terms of trade shock that serves as punishment.

Domestic importers take exports x from their countrymen at the price of the consumption good and

exchange them for m , taking their price e as given. If the country is being punished for default, it receives fraction $1 - \pi$ fewer imports. Their problem is summarized by **Importer's Problem**:

$$\begin{aligned}\Pi(x, m, h) &= em - x, \\ \text{s.t.} \\ m &= (1 - \pi)^h f(x).\end{aligned}$$

Notice that the trade channel penalty, $1 - \pi$ enters as if there were a sudden increase in iceberg costs. This is a conscious choice, intended to expose defaulters to trade barriers that increase the cost of sending their goods abroad. For example, country's exporters may find that its exports sit in customs longer after default, that foreign trade inspectors are more deliberate or, more benignly, that favorable trade agreements are canceled; this can be an efficient way of adding real costs of exporting, and Hummels (2001) shows how more time en route can be a trade barrier. Alternatively, one can think of country getting π less imports for the same amount of exports as a crude defaulted debt renegotiation process.

Figure 2.2 depicts the effect of the trade penalty, given the concave functional form of $f(x)$. The export quantity is a distance from disposable income to consumption of home produced goods. The endogenous exchange rate, e is the slope of the tangent.

2.2.4 International Financial Markets

Denote the default decision control $h(y, b) = \mathbb{I}(W(y) > V(y, b))$. It is equal to 1 when country announces default. World financial markets are risk-neutral but cannot enforce their debt contracts. They have perfect knowledge of the sovereign's problem, so they have zero profits and set $q(y, b')$ so that the expected return equals the international risk-free rate of return R . The credit market **Zero Profit Condition** is:

$$q(y, b') = \frac{1 - E[h(y', b')]}{1 + R}.$$

2.2.5 Recursive Equilibrium

The Recursive Competitive Equilibrium is a collection of

- consumer choice functions $(c_V(y, b), b'_V(y, b), m_V(y, b))$ when the country is not in default,
- consumer choice functions $(c_W(y, b), m_W(y, b))$ when the country is in default,

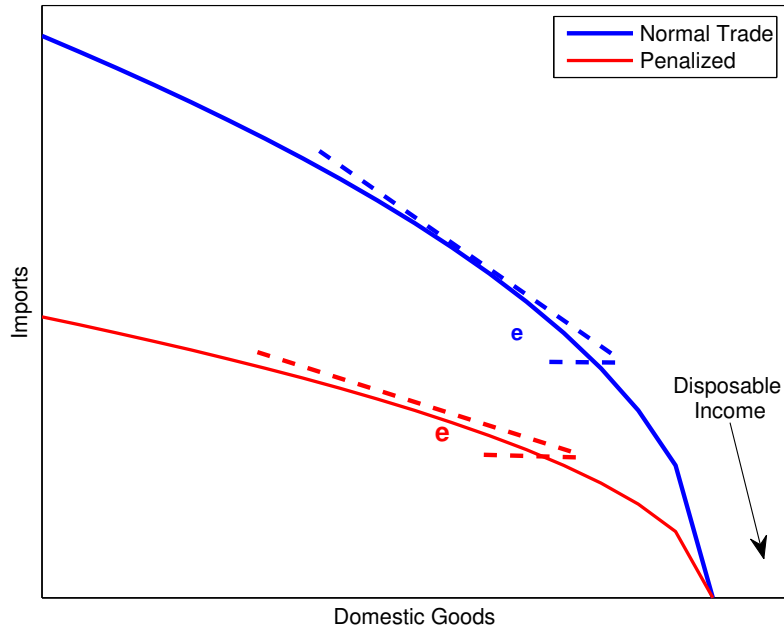


Figure 2.2: The export-import transformation function and the effect of the trade channel penalty

- consumer's default choice function $h(y, b)$,
- consumer's value functions $(V(y, b), W(y), U(y, b))$,
- importer choice variables in no default state $(x_{Vm}(y, b), m_{Vm}(y, b), \Pi_V(y, b))$
- importer choice variables in the default state $(x_{Wm}(y), m_{Wm}(y), \Pi_W(y))$,
- and price variables $(e_V(y, b), e_W(y), q(y, b'))$

such that:

- $(c_V(y, b), b'_V(y, b), m_V(y, b))$ solve the **Household's Problem In Default**, given $\Pi_V(y, b)$, $e_V(y, b)$ and $U(y, b)$, and $V(y, b)$ is the value function of this problem.
- $(c_W(y, b), m_W(y, b))$ solve the **Household's Problem In Good Standing**, conditional on $\Pi_W(y)$, $e_W(y)$ and $U(y, b)$, and $W(y)$ is the value function of this problem.
- $(h(y, b))$ solve the Household Problem conditional on $V(y, b)$ and $W(y)$, and $U(y, b)$ is the value function of this problem.
- $(x_{Vm}(y, b), m_{Vm}(y, b))$ solve the Importer's Problem conditional on $e_V(y, b)$, and $\Pi_V(y, b)$ is the value function of this problem.

- $(x_{Wm}(y), m_{Wm}(y))$ solve the Importer's Problem conditional on $e_W(y)$, and $\Pi_W(y)$ is the value function of this problem.
- $q(y, b')$ satisfies the Zero Profit Condition given on $h(y, b)$.
- $e_V(y, b)$ is such that import market clearing condition $m_V(y, b) = m_{Vm}(y, b)$ holds.
- $e_W(y)$ is such that import market clearing condition $m_W(y) = m_{Wm}(y)$ holds.

Equilibrium exists for the same reason as in Arellano (2008); our problem is separable between borrowing and consumption. After the value of borrowing for the next period is chosen, the allocation of available income between consumption and imports is maximizing monotone function upon compact set.

2.3 Quantitative Evaluations

To evaluate the predictive power of our model, we calibrate our baseline version to Argentine's historical data, following much other related research. Argentine experienced three international default occurrences, in 1982, 1989 and 2002. None of the restructuring periods were particularly long. All these defaults were accompanied by nominal exchange rate depreciation and non-zero trade balances.

2.3.1 Quarterly Data

Our parameter values come from our own estimates and Arellano (2008). Specifically, we estimated the parameters of income time series, the goods relative preference parameters, and import-export relationship parameters and borrowed the values of R , β , ϕ and σ . INDEC, National Institute of Statistics and Censuses, provides quarterly estimates of GDP composites, deseasonalized and in same-year prices, for years of 1993-2008. To make per capita values, we divide by the annual population of Argentina, obtained from CIA Factbook. To convert the import data from peso expenditure into quantities of foreign goods we used the real exchange rate taken from the European Central Bank website².

The parameters that we used were estimated on the quarterly dataset from 1993 to 2008, and are provided in a following table.

Most of our estimates seem close to similar estimations by others. In the international trade literature, most parameter estimates seem to be vigorously debated, though Ruhl (2003) nicely situates the discussion on the proper elasticity parameter κ . Compared to other international business cycle models, our elasticity estimate is slightly high.

²From Statistical Data Warehouse section, located at <http://sdw.ecb.europa.eu/>

Parameter	Value	Explanation	Rationale
R	0.017	Unconditional expected rate of return required by international banking system.	Quarterly return on US 5 year bond
ρ	0.9849	Autocorrelation of log-output.	Estimated
s	0.0258	Standard Deviation of log-output.	Estimated
π	Varying	Trade penalty for defaulting.	
ϕ	0.7180	Probability of being not forgiven on the next period.	Sandleris, Gelos and Sahay (2011)
β	0.953	Subjective time discount factor.	Arellano (2008)
α	0.5859	Parameter of the instantaneous utility function; weight of home good consumption.	Estimated
κ	0.8447	Parameter of the instantaneous utility function; corresponds to 6.441 elasticity of substitution of export to import.	Estimated
σ	2	Parameter of the instantaneous utility function; corresponds to -1 elasticity of intertemporal substitution.	Arellano (2008)
θ	0.2082	Curvature of export-import transformation function.	Estimated
θ_0	0.0467	Relative position of export-import transformation function.	Estimated
θ_1	0.1959	Scale of export-import transformation function.	Estimated

Table 2.2: Parameter values, quarterly data.

For details about the estimation of our import-export parameters, $\theta, \theta_0, \theta_1$, see Appendix B.2.2. We are not aware of others who have estimated a comparable functional. However, Das, Roberts and Tybout (2001) find fixed costs for export supply close to zero using firm-level data. This is consistent with our value for θ_0 . Hummels and Klenow (2005) analyzed the growth of exports from developing countries and found that up to 60% of the increase comes from the “extensive margin.” Rather than deepening existing trade relationships, countries increase exports by diversifying their goods or partners. This seems to be evidence that export markets quickly become satiated and justifies small θ .

2.3.2 Solving for the Recursive Equilibrium

Since competition in the home market is perfect, the solution to the equilibrium is the same as solving with a centralized home economy in which the importer’s problem is internal to the household. We call this the **Centralized Equilibrium**, and describe it in Appendix B.1.

Every period the country faces an import-generating technology, $f(x)$ as in equation 2.4 and allocates its consumption between imports and domestic production, as illustrated on Figure 2.2. Disposable income is $y + b - q(b'(y, b))b'(y, b)$ if country can borrow and y if the country cannot.

To numerically solve this problem, we use value function iteration to solve each subproblem of the Centralized Equilibrium, which is equivalent to solving The Household Problem In Default and The Household

Problem In Good Standing. Each of these value functions converges trivially. The Household Problem is characterized by the max of these two value functions, and default occurs where W exceeds V . Because both functions are monotone in y and $-b$, for a given value of b , they may cross. As discussed in Arellano (2008), we should have a single crossing, which means that there is a single $b(y)$ that defines the maximum level of debt before the country defaults, given a realization of y .

In the Argentine default example, the estimate of $\ln(1 - \pi)$ is -0.6939, which implies $\pi \approx 0.5$. We take this estimate as illustrative as there is no international policy for punishing defaulters, so there is no guarantee that one estimate for π should hold in other contexts. We evaluate the model's predictions at various punishment levels, to obtain bounds on plausible outcomes, and to account for potential short-sample biases. Varying the value of $\pi \in \{0.2, 0.5, 0.8\}$ also serves to demonstrate the effect of the trade channel punishment. We see that the country's default policies and behavior in default is sensitive to the magnitude of the trade channel punishment.

2.3.3 Welfare Comparisons

To understand our model with incomplete markets, we compare its agents' welfare to two extremes: complete asset markets and absent asset markets. With incomplete markets, bankruptcy provides some insurance against uninsurable, idiosyncratic risk, as discussed in Livshits, MacGee and Tertilt (2007). π determines the cost of such insurance, and affects both consumption and expected utility. For comparisons, we set $\pi = 0.5$ and other parameters follow Table 2.2. The expected value function in stationary distribution is -18.4318 and the average volume of the consumption aggregate, $(\alpha c^\kappa + (1 - \alpha)m^\kappa)^{1/\kappa}$, is 0.5868. The certainty equivalent of aggregate consumption is 0.5358. The risk premium is 9%, defined as $\frac{E\tilde{C}}{c_C E} - 1$ where \tilde{C} is the actual aggregate consumption and $c_C E$ is an aggregate consumption certainty equivalent.

	Expected utility	Aggregate consumption CE	Risk premium	CE difference w/ benchmark
Benchmark	-18.4318	0.5358	9%	
Financial Autarky	-19.7300	0.5189	3%	-4%
First Best	-15.3818	0.5804	0%	+8%

Table 2.3: Welfare comparison.

The first best scenario provides agents with complete asset markets, which implies perfect consumption smoothing. At the other extreme, in financial autarky, the country cannot save or borrow and cannot insure against income shocks. We elaborate on the results in Appendix B.3 and they are summarized in Table 2.3.

2.3.4 Equilibrium Results

We obtain three interesting equilibrium predictions. First, country borrows more if it had a history of negative income shocks, i.e. high current borrowing. Second, the current account is countercyclical, in line with data and Aguiar and Gopinath (2006). Third, positive shocks to income depreciate the home currency. The borrowing policy result is immediate, as it is a consequence of risk-aversity of a representative consumer. Most endogenous default papers have this result.

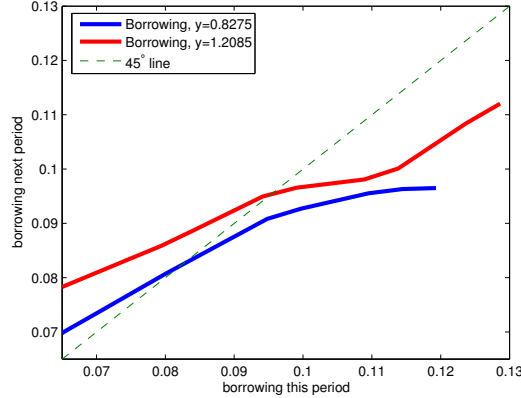


Figure 2.3: Borrowing policy. Curves stop on right at the point of default.

Our countercyclical current account (Figure 2.4) is consistent with data and, as Aguiar and Gopinath (2006) point out, nontrivial. In a partial equilibrium, positive shocks would induce savings and a positive current account, but if the interest rate is endogenous, it falls with positive income shocks (because the probability of default is counter cyclical) and spurs borrowing. Arellano (2008) has a pro-cyclical current account, and Aguiar and Gopinath (2006) shows that in their framework, without demand for trade, the current account's correlation depends on the specification of income process. In our model, the current account is governed by both consumer preferences for foreign goods, the importing technology and consumption smoothing. In this richer trade environment, the behavior of the current account does not just mirror the agents' borrowing policy. As our major contribution relies on the trading process that we model, we treat this result as a sign that our trading process properly approximates reality.

The counter cyclical exchange rate is a new result in the endogenous default literature, though Chari, Kehoe and McGrattan (2002) and others establish a similar result in data and in various models. Chari, Kehoe and McGrattan (2002) shows that a fairly standard RBC can generate counter cyclical real exchange rate movements. The intuition within our model is that with an increase in y , the country's possibility frontier in terms of (c, m) moves outwards. Consumption of both goods increase because the goods are

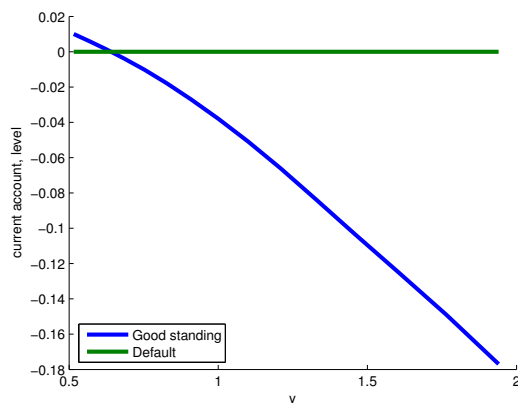


Figure 2.4: Current account conditional on output in stationary distribution

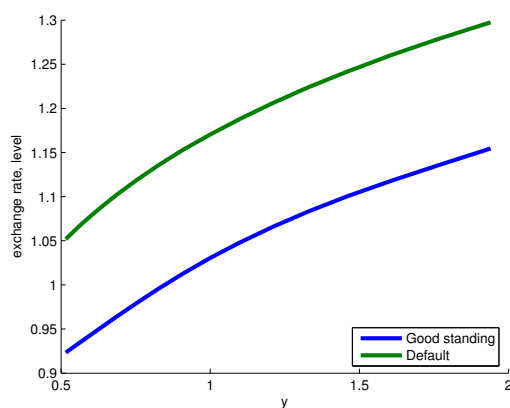


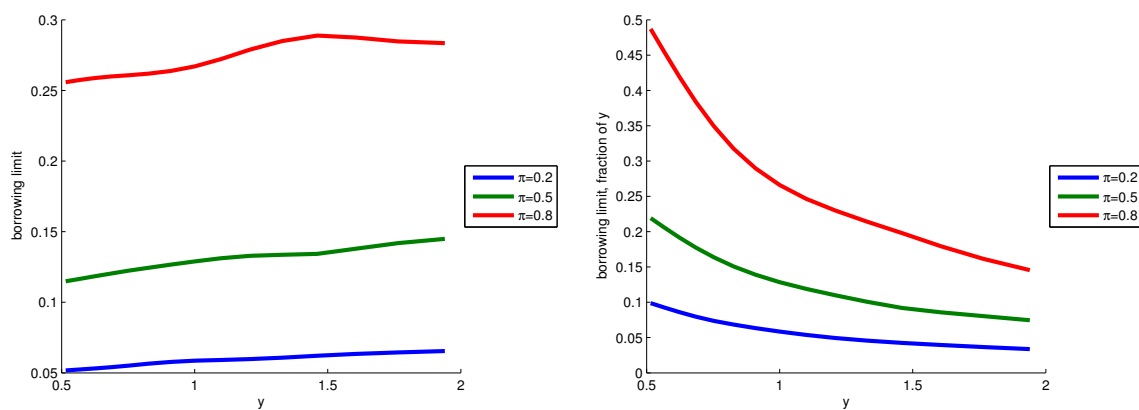
Figure 2.5: Exchange rate conditional on output in stationary distribution

normal. Since the marginal utility of consumers abroad goes down with additional exports, the exchange rate must rise with the increase in demand for foreign goods. It is important, however, to understand that foreign consumers in the model have no income shocks. To compare our result to the data, one must consider potential shifts in the world’s demand function, which would manifest itself in Equation (2.4) and may shift the tangency point in Figure 2.2.

Notice also that the exchange rate is higher in default, though not $\frac{1}{1-\pi} = 2$ times higher — that happens because of the change in trade volumes due to substituting away from imports. The penalty shifts the line down, but the households are also able to adjust their consumption basket, and due to the curvature of Equation 2.4, this affects the exchange rate. As discussed earlier, we could generate a depreciation with linear trade technology, but the nonlinearity due to $\theta < 1$ allows for this additional freedom. As the data reveal, there is variation in the size of depreciation and trade volumes tend to shift, both consistent with our preferred parameterization.

2.3.5 The Effect Of Penalty Levels

Figure 2.6 depicts the default decision threshold, the debt such that the country defaults if borrowing goes any higher. The small penalty case, $\pi = 0.2$, in which very little debt can be supported recalls the result of Bulow and Rogoff (1989). As in Aguiar and Gopinath (2006) and Arellano (2008), the sustainable debt increases in y ; which makes the interest rate spread counter cyclical. The right panel of Figure 2.6, demonstrates, that the borrowing limit does not increase as fast as y , so that an expansion does not imply that a borrower can support a higher percentage of debt. This monotone slope, however, reflects the particular parameterization. With different specification (namely, low θ), it can yield an interestingly U-shaped borrowing limit. With this result, recessionary countries with low y , have greater incentive to stay in good standing because they expect the negative shock to persist; and countries with high output realizations are also less likely to default; however, in the middle, countries default on much smaller debt level.

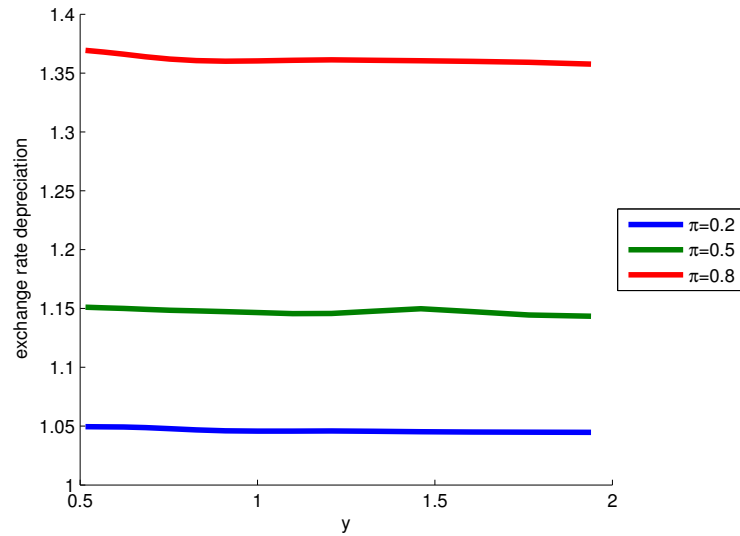


Note: Argentine had debt of around 30% of GDP at the end of 2001, and normalized detrended GDP of 0.95.

Figure 2.6: Default decision borrowing threshold

The penalty value changes the exchange rate adjustment post default, as seen in Figure 2.7. The exchange rate change pictured is calculated as the difference between the exchange rate in a default state and the exchange rate just before the default. That is, if the country has a debt that leaves it indifferent between defaulting and remaining in good standing, what would be the difference of exchange rates in these two cases? This thought experiment about the moment of default is close to thinking about two otherwise equal countries, one slightly below the default threshold, and one just above. Such a comparison isolates the change purely as a response to default decision. The figure reveals that the level of penalty determines the size of the depreciation. However, the country moderates some welfare damage by substituting domestic product for import consumption. The amount of depreciation does not depend much on the country's position in the business cycle.

Figure 2.8 (left) illustrates the substitution behavior of domestic households. Country seem to have the



1.15 on this figure means that foreign goods cost 15% more after default than before default. Penn World Tables report 18% change of PPP index between 2001 and 2002.

Figure 2.7: Exchange rate drop due to default

same consumption policy at the moment of default independently of the level of π , that's why consumption policy before default is represented by a single line (this qualitatively holds for other parameterizations). As the trade channel penalty rises, the substitution effect becomes stronger. Figure 2.2 reveals the link between the change in the makeup of the consumption basket and the exchange rate that we highlighted earlier. Home good consumption is a normal good, and in a default state, an increase in income leads to an increase in consumption; while in the good standing state, higher income allows the country to borrow more, somewhat negating the decreasing returns to scale of import-export relation. A positive shock to income in a defaulted country will lead to increased consumption of home goods; a positive shock in non-defaulted country can lead to a more than one-for-one increase in exports.

Figure 2.8 (right) demonstrates the pattern of import's share in consumption. Decreasing returns to scale of the export-import relation make the dependence on y negative: countries in expansion export ever greater amounts to get additional units of imports. Trade channel penalties have both direct effects on the level of import (by construction of the model), and indirect effects through substitution. As in the data, proportional expenditure on imports falls following default; countries in expansion after default benefit from not servicing the debt (which they'd have to if they did not defaulted), which instead they can spend on consumption, including imports. Notice that share of imports for high y countries exceeds $(1 - \pi)$ times the corresponding value before default.

Figure 2.9 shows the change in capital account and trade balance at the moment of default decision. The

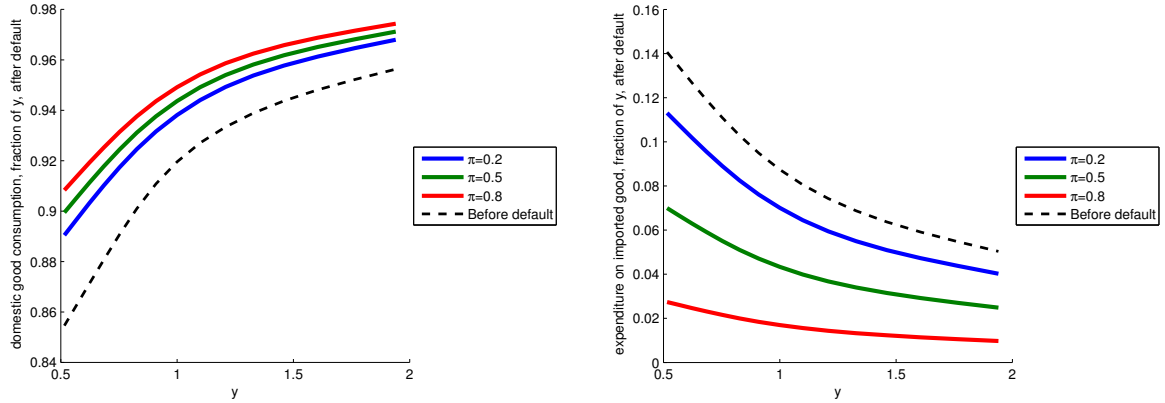


Figure 2.8: Domestic good consumption (left) and foreign good expenditure (right)

trade balance improves more drastically in countries in downturn and the penalty level contributes significantly to the change size. The difference between the change in trade balance and change in capital account is the change in assets inflow. This might involve foreign currency reserves of respective central bank. Particularly, it implies that a country that defaults in a cyclical expansion and faces $\pi = 0.8$ would have to come up with additional assets equal up to 5% of GDP. Combined with a policy of fixed exchange rate and capital controls, that extra inflow can create a greater than necessary nominal money mass, resulting in inflation.

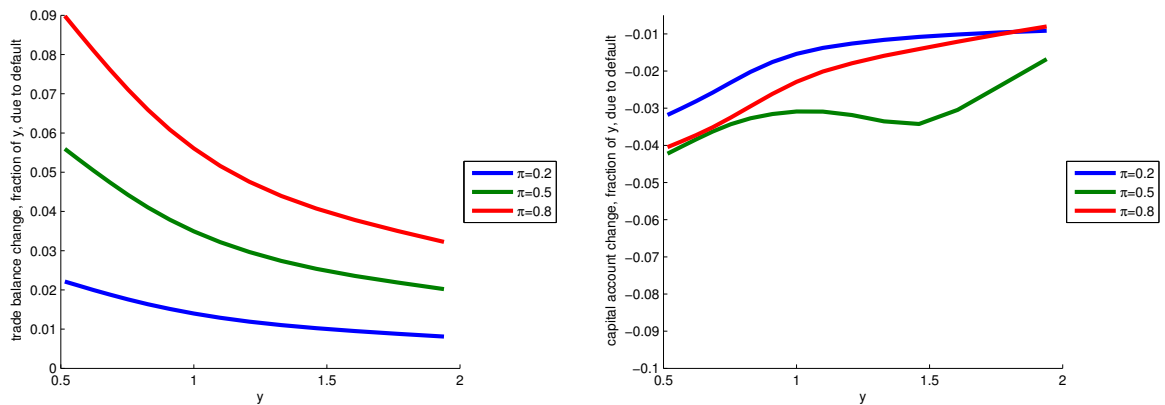


Figure 2.9: Trade balance (left) and capital account (right) changes at default.

To summarize, a higher trade channel penalty leads to: (a) greater household adjustment through the domestic consumption basket, and trade quantities, (b) higher sustainable debt, and (c) greater exchange rate depreciation.

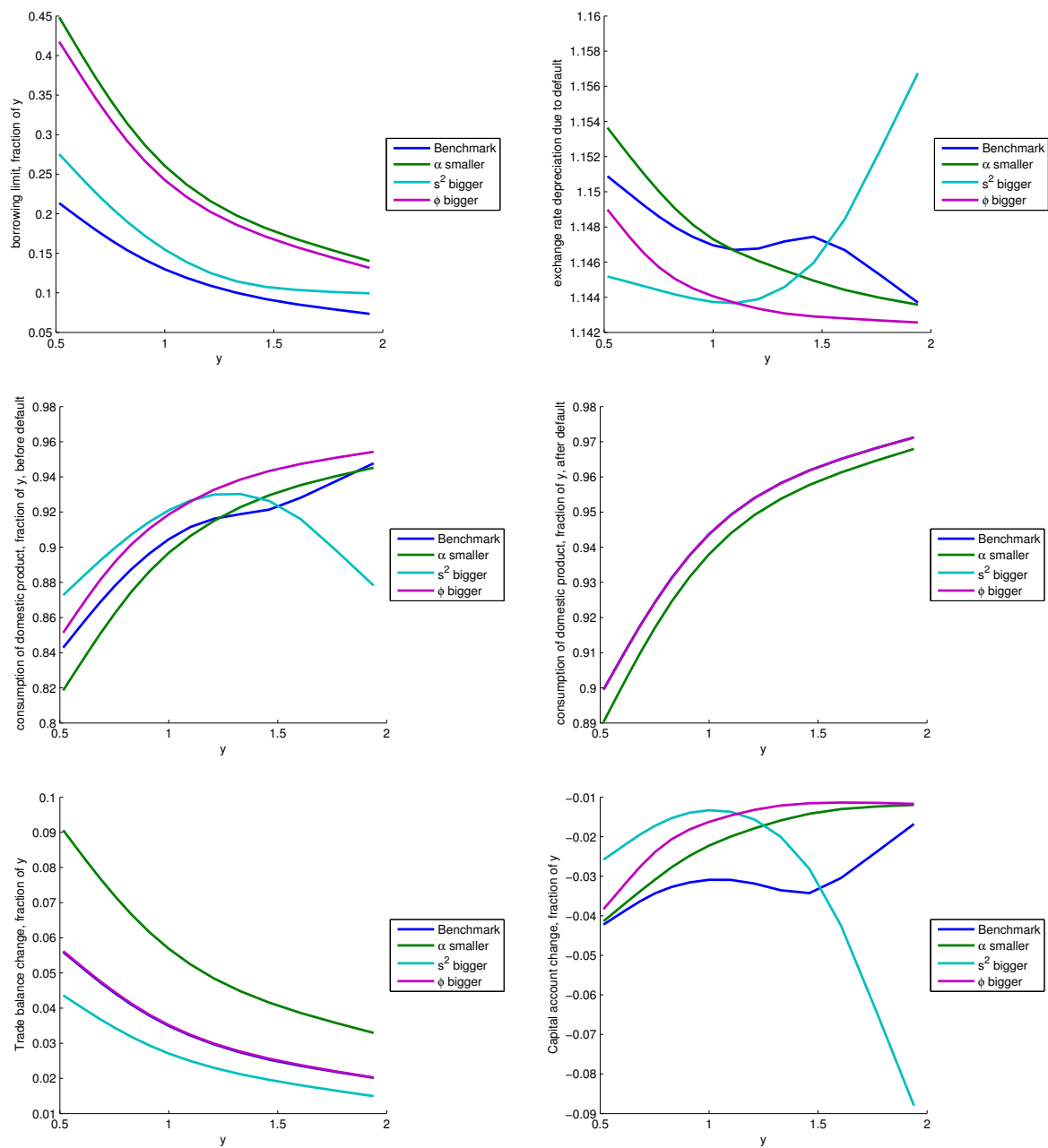


Figure 2.10: Comparative statics.

2.3.6 Comparative Statics

Figure 2.10 summarizes the expected differences that should arise in cross-country data. The parameters we consider are a change in preferences for imports (α decreased by 0.1), a change in ability to forgive (ϕ increased by 0.1) and a change in variance of income (s^2 doubled the value). The penalty π in benchmark case is set at 0.5.

Changing α affects the willingness to default. A country that is less import-oriented defaults sooner and consumes less import in all states. A rapid change in tastes toward home consumption could precipitate default. Consider a government with debt just below the maximum allowable with a certain value for α ; then, if this sovereign suddenly cared less about the import component of consumption, the level of allowable debt will shift down, possibly below the existing stock. In a populist coup, the new government might internalize the preferences of a different social group, one which consumes relatively fewer imports. This abrupt change of tastes of the sovereign's constituency — if combined with severe pre-existing debt — could result in default and explain some of the “inexcusable” defaults of Tomz (2007).

Loading this experiment on α with so much meaning is knowingly problematic. It is not actually a story consistent with our model, as our CES preferences are homothetic, so a poorer social strata should still have the same demand for imports. Further, our country's household is representative. If a coup implies a change to the preferences that the sovereign internalizes we should have to justify this with some explicit strategy for aggregation a diverse population. In the simplest case, our result is equivalent to a model in which a sovereign that puts full weight on his median constituent. Then, if this groups changes, the preferences upon which he acts would also change.

After default, there is a stronger improvement in trade balance when α is lower, because a lower α country will forego more home consumption in both relative and absolute respects. The new preference regime, by consuming fewer imports, moves down the decreasing returns import-export technology, and so in the aftermath of a default, can sell into a “steeper” market.

The variance of the income process does not notably affect the patterns of trade adjustment relation, either before or after default. However, it affects the default decision through the financial exclusion penalty. This does not necessarily conflict with the Bulow and Rogoff (1989) finding that financial exclusion cannot sustain debt, because we also exclude countries from saving while in bankruptcy. Countries in expansion are more motivated to sustain debt than medium-range countries. The intuition is that high income variance countries obtain higher benefits from maintaining good standing in order to be able to smooth consumption in the future.

Also, observe that the variance of the composition of import and home goods is higher when y is high,

and therefore the impact of increased variance is stronger on the right tail of income distribution. That leads to greater aversion to financial autarky, and consequently, to bigger tolerance of debt. Therefore, countries with higher income volatility might actually be safer for investors than steady economies.

The value of ϕ affects the cost of default, but once the choice between defaulting and staying prudent has been made, ϕ does not affect other tradeoffs, and therefore does not change the policies significantly. The bigger is ϕ , the worse is a country's cost of default and this its aversion becomes even stronger if it is currently experiencing an expansion. However, bigger ϕ demands more defiance from international lenders: longer punishment means loss of positive NPV projects and restraining selves from exports of the punished country. These concerns over enforcement are explored more in Wright (2001), but we just take as given that penalties are credible.

The comparative statics demonstrate that our model is qualitatively robust to parameter values, and provide additional insights into parallel problems (such as the consequences of populist uprisings and international risk management). The existing literature noted that business cycle fluctuations do not account for a significant fraction of historical default episodes. Our comparative statics demonstrate how changes in underlying parameters immediately result in significantly different equilibrium strategies — and confuse the identification of the business cycle contribution. Moreover, we can see that default boundary condition on debt shape differently with respect to output fluctuations with respect to these parameters; countries will demonstrate different default behavior. Even “inexcusable” in Tomz (2007) sense behavior is not outside of offered model's scope; it only implies certain heterogeneity of parameters.

The top right plot of Figure 2.10 demonstrates that changing parameters does not change equilibrium depreciation much (and parameters that are changed, are changed a lot). In fact, only two parameters seem to matter for depreciation: κ and π . Other parameters, chosen to match Arellano (2008) do not affect equilibrium default depreciation, which seems to be arising from the combination of CES utility function and the proportional penalty. With respect to other parameters, our model is (knowingly) ill-suited to match certain moments of the data. For example, without long-term contracts Chatterjee and Eyigungor (2009) suggests models will not be able to match default frequency without undue manipulation of other parameters. In the interest of conservatism, we kept many parameters from the literature. Equilibrium depreciation, our chief concern, depends mostly upon the few parameters that we have discussed.

2.4 Conclusion

Our paper introduces a model that accounts for the systematic linkage between sovereign default and real exchange rate depreciation. Countries interact with the rest of the world for two reasons: they borrow to smooth consumption over the business cycle and trade for goods that cannot be produced at home. They may default on their borrowing, but misbehaving on the international stage hurts their international trade. A country in default has to export away more units of its own production for the same amount of imports. Consumers substitute away from the foreign good, and the new exchange rate, the equilibrium price of foreign goods, depreciates. Because trade disruption is costly in terms of welfare, it reduces a country's willingness to default and allows it to sustain positive debt.

Prior research has noted this empirical connection, but has not explicitly modeled it. Other quantitative models capture the contribution of output fluctuations to the default strategy; however, they consider international markets only as imperfect insurance against stochastic income shocks, abstracting from the desire for differentiated goods from abroad. Here, we have highlighted the importance of international trade. Our trade channel penalty allows a country to support a realistic level of debt before default. Furthermore, the country's maximum borrowing limit is not necessarily increasing with income, a feature observed in the data but difficult to match with other models. We explore the model's response to changes in determinants such as income volatility and penalty duration. The qualitative results of these parameter manipulations are generally consistent with common intuition. We investigate their effects on default and trade policy, and generate policy recommendations for lenders.

In this exercise, we calibrated to Argentina to demonstrate the predictions of our import-export mechanism, which is the real innovation of our paper. The parameters obtained from this calibration generated reasonable predictions for the default decision and subsequent international adjustments. Future research might try to match other default episodes more closely. From this basic model, we can also integrate nominal fluctuations and financial flows to capture inflation and credit crunches. This paper, however, elides nominal concerns to emphasize and clarify the real factors behind the default-linked depreciation.

We begin the discussion about how international trade parameter changes could change the incentives to default. It seems that more import-dependent countries default less frequently. Future work should more formally consider a mixed-model of default. A particularly promising avenue is to integrate parameter risk into endogenous variables, e.g. how does the probability of a coup effecting preferences for abroad goods modify the interest rate's default premium? Our discussion also raises a new potential for models with heterogeneous agents and competing interests. Our simple model, considering trade and default, might be fruitfully combined with various other branches of the study of crises.

Chapter 3

University Competition, Grading Standards and Grade Inflation

3.1 Introduction

Universities award grades to measure the performance of students in courses. In turn, important decisions by third parties are based in part on GPAs — firms tend to offer higher wages to students with high GPAs, and graduate schools tend to admit high GPA students¹. In this paper, we study how universities choose grading standards when they care about the decisions made by third parties based on GPAs. We characterize how student body qualities at different schools interact with the depth of the job market to affect equilibrium grading standards.

Our model reconciles three central empirical regularities describing grading over the past fifty years: (1) GPAs are higher at better schools, (2) GPAs have risen over time at all schools, and (3) grading standards have fallen faster over time at better schools. It is manifest that better universities award more high grades. For example, Rojstaczer (2003) finds that GPAs at private universities in 2006-2007 are 0.3 higher than at public universities. Table 3.1 reinforces these findings, presenting the evolution of grades at selected universities between 1960 and 2000. This table reveals that grades at better universities are uniformly higher. The table also highlights a uniform secular rise in GPAs over time. In addition, over the entire sample period, GPAs at better universities increased significantly faster, although there is no significant difference in grade inflation in different universities between 1980-2000.

We develop a model in which universities are distinguished by the distributions of “academic abilities” of their students: the distribution of student academic abilities at top schools conditionally stochastically dominates that at lesser schools. Firms value both “academic ability” and social skills, which are complements in production. There are two types of jobs, good and bad, which are distinguished by the higher marginal product of skills in good jobs. Good jobs are in limited supply. Universities determine which students receive “A” grades by setting endogenously-chosen cutoffs on academic ability. Firms learn student social skills via job interviews, and forecast academic abilities using the information contained in the ability distribution

¹There is a large empirical grading standard literature; see Bar and Zussman (2011), Rose and Betts (2004) and Bagues, Labini and Zinovyeva (2008) for both the questions they study and their literature reviews.

	1960	1980	2000	1960 - 2000	1980-2000
Harvard U	2.7	3.05	3.41	0.71	0.36
Princeton U	2.83	3.13	3.36	0.53	0.23
Yale U	2.56	3.27	3.48	0.92	0.21
Columbia U		3.2	3.36		0.16
Stanford U	2.98	3.27	3.55	0.57	0.28
Northwestern U		3.02	3.35		0.33
U Chicago	2.5		3.26	0.76	
MIT	2.47	3.27	3.26	0.79	-0.01
Dartmouth C	2.47	3.06	3.33	0.86	0.27
Duke U	2.41	3.02	3.36	0.95	0.34
Average grade inflation				0.7613	0.2411
Standard deviation				0.1531	0.1145
U Illinois	2.77		3.12	0.35	
U Miami		2.7	3.05		0.35
Penn State U		2.86	2.99		13
U Wisconsin	2.51	2.89	3.13	0.62	0.24
U Texas		2.6	3		0.4
U Washington	2.31	2.97	3.12	0.81	0.15
UC Irvine		2.9	2.95		0.05
Lehigh U		2.6	2.97		0.37
Georgia Tech		2.78	2.97		0.19
George Washington U		3.03	3.25		0.22
Average grade inflation				0.5933	0.2563

GPA Data from GradeInflation.com. When year X was not available, the closest year was used; if there was no observation in ± 5 years, the observation was treated as missing. The top list of schools is from the top list of US News National Universities rankings, and the bottom list is from the bottom 50 and up of the same list for the sub-samples that had grade points in at least two time periods. Grade inflation is calculated as difference in GPAs.

Table 3.1: Evolution of GPAs at selected universities.

at a student’s university, the university’s grading standard and the student’s grade. Firms then make job assignments, and wages are competitively determined.

Universities understand how firms determine job placement and wages, and set grade standards to maximize the total expected wages of their graduates. Top universities would argue that their higher proportions of high grades simply reflect their better student bodies; a common grading standard would inevitably lead to more good grades at better schools. A central result of our analysis is that under weak conditions, top universities actually set *softer* grading standards: the marginal “A” student at a top university is *less* able than the marginal “A” student at a lesser university. The intuition for this result devolves from the basic observation that a marginal student at a top school can free ride on the better upper tail of students because firms cannot distinguish “good A” students from “bad A” students. In contrast, lesser schools must compete for better job assignments by raising the average ability of students who receive “A” grades, setting excessively high grading standards. It is the competition for good job assignments for graduates that generates this result—while giving more “A” grades lowers the expected productivity and hence wages of students who receive good job assignments, additional well-placed students more than offset this wage effect in the eyes of university.

Importantly, a social planner who seeks to maximize total output in society would choose the *opposite*

ordering, setting more demanding grading standards at top schools whenever there is heterogeneity among students in social skills. Were academic skill the only source of heterogeneity, the social planner would set the same grading standard at each school. However, then the expected academic productivity of students with “A”s at top schools is higher. Firms, which do not see the abilities of students, would then “over-rate” marginal “A” students from top schools, assigning too many with low social skills to good jobs. The social planner wants to equalize the expected productivity of the marginal “A” students who receive good job assignments, and since the average social skill is less for students from top schools with good job assignments, the social planner sets a higher grading standard for “A”s at top schools.

Finally, we identify a plausible driving force underlying grade inflation at universities: the secular increase over time in the measure of good jobs relative to the measure of students, possibly reflecting the well-established shift toward skill-biased technologies. There is extensive evidence that skill demands at jobs have increased significantly, suggesting that there are now far more good jobs. We show that universities respond to an increase in good jobs by trying to place less able students at good jobs. In particular, universities reduce cutoffs for “A” grades.

The upward trend in grades is perhaps less interesting than the inference problems it creates for third parties. In and of itself, grade inflation might not increase inefficiency in the economy. However, we provide conditions under which grade inflation exacerbates differences in grading standards, further distorting hiring decisions. That is, we provide conditions under which the cutoff for an “A” falls by more at top schools. Intuitively, if there are very few good jobs, grade standards must be very high, so that with a common bounded support on ability, ability differences between the marginal “A” and average “A” student must be small at all schools. However, when more students get “A”s, this difference grows, and top schools exploit this via lowering their grading standards by more. We then provide conditions under which this greater reduction in grading standards at top schools is associated with a greater increase in the number of “A”s, i.e., for grade inflation at top schools to be higher.

We next review the literature. Section 2 presents the basic model of university competition. Section 3 derives equilibrium and social planner outcomes when all students have the same social skill. Section 4 extends the analysis to a setting in which students differ in social skills. Section 5 summarizes our findings and discusses the robustness of our findings to perturbing our assumptions. An appendix contains all proofs. **The Literature.** The paper closest to ours is probably a free-rider paper of Yang and Yip (2003), which also predicts more good grades in better schools. However, this paper predicts that all students receive the same wage, independently of grade: in their equilibrium, grades convey no information about students. Moreover, in their model, universities intentionally destroy value by explicitly lying about students’ abilities.

Chen, Hao and Suen (2007) consider a setting in which the measure of good students at a school is random, observed by schools, but not by firms. They model the intentional loss of academic reliability where the grading standard is not fixed, so that otherwise identical students who take identical actions might receive different grades. They argue that this is why grading standards have varied over time. However, they are silent as to why there should be significant unobservable year-to-year variation in the quality of large populations of students at a university, especially given indirect, but broadly observable measures of student quality such as mean SAT and ACT scores, and measures of high school class rank.

Dubey and Geanakoplos (2010) investigate how discreteness of grades influences student effort: they find that when students only care about relative rank, coarser grade structures can motivate students to study harder. Without noise in assessment ability, there is little reason to have a big support for grades, at least at the right tail. Ostrovsky and Schwarz (2010) study the optimal transcript structure problem, and find reasons to give the same grade to a sizeable fraction of students in the right tail, when many schools have worse ability distributions. Zubrickas (2010) studies the optimal mechanism for eliciting effort from students by awarding grades, and finds that with imperfect ability assessment, to elicit effort from the best students there must be a positive mass of students with the best grade. MacLeod and Urquiola (2009) explore how the structure of the schooling market affects tradeoffs between studying effort, wealth and leisure.

A body of literature studies grading standards from the perspective of a central planner. Costrell (1994) studies how different policies toward standards affect student effort (he states that an egalitarian central planner is likely to pick lower grading standards than a total earnings maximiser), and provides a review of the grading literature; Betts (1998) provides an opposing argument.

3.2 The Competition Between Universities

The world contains two types of *universities*, $u \in \{H, I\}$. Universities are distinguished by the ability distributions of their student bodies. Abilities at a type H school are distributed according to a density $f_H(\theta)$, and the distribution at a type I school is $f_I(\theta)$. These densities are continuous and strictly positive on their common support, $[\underline{\theta}, \bar{\theta}]$. We capture the notion that the student body at a type H school is better with the concept of conditional first-order stochastic dominance: $f_H(x|x > t)$ first order stochastically dominates² $f_I(\theta|\theta > t)$ for all $t \in [\underline{\theta}, \bar{\theta})$, written $f_H(\theta) \succeq_C f_I(\theta)$. In particular, the associated cumulative distribution functions satisfy $F_I(\theta|\theta > t) > F_H(\theta|\theta > t)$, for all $t \in (\underline{\theta}, \bar{\theta})$ and $\theta \in (t, \bar{\theta})$.³ To capture the fact that any

²This is equivalent to a hazard ordering $\frac{f_H(x)}{1-F_H(x)} < \frac{f_I(x)}{1-F_I(x)} \forall x$ and is implied by likelihood ordering: $\left(\frac{f_H(x)}{f_I(x)}\right)' > 0$.

³As we show later, all of our findings extend directly and immediately to arbitrary numbers of university types, provided that the distribution of student body academic abilities at different university types is ordered with respect to conditional first-order stochastic dominance.

single university admits a negligible portion of the entire pool of students, we assume there is a continuum of each type of university. The total measure of students is normalized to one, and measure $\alpha \in (0, 1)$ of students attend type H universities.

A student is distinguished by his (a) university type, (b) academic ability, θ , and (c) social skill, μ . Social skills, μ , are distributed according to the density $g(\cdot)$ and distribution $G(\cdot)$ with nonnegative full interval support $[\underline{\mu}, \bar{\mu}]$, and are distributed independently of academic skills. We assume that the distribution of social skills is the same at all schools. This assumption reflects the observation that universities largely filter students via high school academic performance and academic tests such as the SAT or ACT. We make the standard increasing hazard rate assumption on g and f_u .

Both academic ability and social skills contribute to the work productivity of a student. Via job interviews, firms can observe μ , but they do not directly observe θ . There are two types of jobs. There is a positive measure Γ of “good” jobs in which the product of a student with ability (θ, μ) is $S\theta\mu$, and many “bad” jobs in which the product is $s\theta\mu$, where $S > s > 0$. Our qualitative findings extend to the class of technologies in which a worker’s output is proportional to $\theta^\alpha\mu$ for some $\alpha > 0$, reflecting that we do not impose strong structure on the distributions of θ and μ .

Universities know the academic abilities of their students, but not their social skills.⁴ Their problem is to assign a grade $g \in \{A, B\}$ to each student.⁵ Universities seek to maximize the expected sum of wages earned by graduates.

Firms do not observe student academic abilities. However, firms know the alma mater of each student and the distribution of academic abilities at each school, and hence can extract information about academic abilities from grades. We will show that in equilibrium students who are expected to have the highest productivity will be assigned to good jobs.

We assume that universities adopt grading strategies that take the form of a cutoff, so university u gives a student an “A” if and only if his academic ability θ exceeds a cutoff $\hat{\theta}_u$ chosen by the university. The same equilibrium outcome would obtain were the labels “A” and “B” reversed: we adopt the convention that an “A” grade refers to the better subpopulation of students. In addition, since giving all students “A” grades is equivalent to giving all students “B” grades, without loss of generality, we assume that if it is optimal for a university to give all students the same grade, then it gives all students “A” grades, as at Doonesbury’s fictional Walden University.

⁴Equivalently, universities could treasure academic integrity so that only θ affects grades.

⁵Albeit this is a strong assumption, we use it to proxy for the idea that there is little difference between 3.8 and 3.9 GPA students, but a lot of difference between 3.9 and 2.9 GPA students. Lizzeri (1999) argues why universities are not interested in revealing too much information. Dubey and Geanakoplos (2010) suggest a story for why a coarse signal structure might help student motivation. In fact, many graduate programs formulate admission requirements in the form of thresholds, and these thresholds are largely consistent among departments.

More generally, our model is sufficiently sparse that equilibria can exist in which grading strategies do not take a cutoff form. Non-cutoff strategies can emerge in equilibrium simply because once firms form beliefs (which determine job assignments), they are not affected by which set of student abilities receive “A”s. However, such equilibria are not robust to natural refinements that pin down what universities do: non-cutoff strategies cannot be part of an equilibrium if either (a) firms observe the true ability of a small measure of students (and schools do not know which ones), from which firms infer average productivities of “A” and “B” students, or (b) employment continues for two periods, and firms learn a worker’s true ability after the first period, and there is either complementary learning-by-doing or workers cannot be reassigned. Under such scenarios, universities have strict incentives to ensure their most able students receive “A”s, and hence that equilibrium grading strategies take cutoff forms.

We denote a student from a type u university with grade g as a ug student. The zero profit condition for firms implies that a student who is hired for a good job receives wage $SE[\theta|g, u, \hat{\theta}_u]\mu$, while a student with a bad job earns wage $sE[\theta|g, u, \hat{\theta}_u]\mu$.⁶ The expected ability of a ug student is $E_{ug}[\theta] = \frac{\int_{\underline{\theta}}^{\bar{\theta}} \mathbf{I}(\text{grade is } g)\theta f_u(\theta)d\theta}{\int_{\underline{\theta}}^{\bar{\theta}} \mathbf{I}(\text{grade is } g)f_u(\theta)d\theta}$.⁷ Notice that issuing fewer “A”s raises the expected academic ability of *both* “A” and “B” students: increasing the grading standard $\hat{\theta}_u$ increases both $E_{uA}[\theta]$ and $E_{uB}[\theta]$. Also, were universities to set a common grading standard, then a type H university would have more “A” students because $f_H(\theta) \succeq_C f_I(\theta)$.

Firms make job assignments to maximize profits. The complementarities between job quality and student productivity imply that in equilibrium it will be students with higher expected productivities who are assigned to good jobs. That is, firms will assign a student from university u with grade g and social skills μ to a good job if and only if the student’s expected productivity $\mu E_{ug}[\theta]$ exceeds a critical endogenous equilibrium standard, K . That is, K denotes the lowest expected productivity among students employed on good jobs, and $K/E[\theta|u, g, \hat{\theta}_u] \equiv \hat{\mu}_{ug}$ is the minimum level of social skills required from a ug student for placement at a good job. Each university is too small to affect the productivity standard K , but each university internalizes how its grading standard $\hat{\theta}_u$ affects the hiring standard $\hat{\mu}_{ug}$ set by firms. We assume that firms set wages equal to expected productivity of hired students.

Given an equilibrium productivity standard $K \leq \bar{\theta}\bar{\mu}$ for a good job, a type u university maximizes the

⁶Much of the literature (e.g., Yang and Yip (2003) and Coate and Loury (1993)), imposes the restriction that everyone at the same job earns the same wage, regardless of their expected productivity.

⁷If the denominator is 0, we set $E_{uB}[\theta] = \underline{\theta}$ and $E_{uA}[\theta] = \bar{\theta}$ to preserve continuity; $\mathbf{I}(\cdot)$ is the indicator function.

total income of its student body by choosing $\hat{\theta}_u$ to maximize

$$\begin{aligned} \pi_u = & \max_{\hat{\theta} \in [\underline{\theta}, \bar{\theta}]} S \int_{\hat{\mu}_{uA}}^{\bar{\mu}} \int_{\hat{\theta}_u}^{\bar{\theta}} \mu \theta dF_u(\theta) dG(\mu) + s \int_{\underline{\mu}}^{\hat{\mu}_{uA}} \int_{\hat{\theta}_u}^{\bar{\theta}} \mu \theta dF_u(\theta) dG(\mu) \\ & + S \int_{\hat{\mu}_{uB}}^{\bar{\mu}} \int_{\underline{\theta}}^{\hat{\theta}_u} \mu \theta dF_u(\theta) dG(\mu) + s \int_{\underline{\mu}}^{\hat{\mu}_{uB}} \int_{\underline{\theta}}^{\hat{\theta}_u} \mu \theta dF_u(\theta) dG(\mu) \\ \text{s.t. } & \hat{\mu}_{uA} \frac{\int_{\hat{\theta}_u}^{\bar{\theta}} \theta dF_u(\theta)}{\int_{\hat{\theta}_u}^{\bar{\theta}} dF_u(\theta)} = \hat{\mu}_{uB} \frac{\int_{\underline{\theta}}^{\hat{\theta}_u} \theta dF_u(\theta)}{\int_{\underline{\theta}}^{\hat{\theta}_u} dF_u(\theta)} \geq K. \end{aligned}$$

Subtracting the total productivity of all students of university u were they all employed on bad jobs (a constant that does not depend on the grading standard) from π_u , we can rewrite the university's objective as

$$(S - s) \int_{\hat{\mu}_{uA}}^{\bar{\mu}} \int_{\hat{\theta}_u}^{\bar{\theta}} \mu \theta dF_u(\theta) dG(\mu) + (S - s) \int_{\hat{\mu}_{uB}}^{\bar{\mu}} \int_{\underline{\theta}}^{\hat{\theta}_u} \mu \theta dF_u(\theta) dG(\mu).$$

Dividing this result by $S - s > 0$ yields:

$$\Pi_u = \max_{\hat{\theta} \in [\underline{\theta}, \bar{\theta}]} \int_{\hat{\mu}_{uA}}^{\bar{\mu}} \int_{\hat{\theta}_u}^{\bar{\theta}} \mu \theta dF_u(\theta) dG(\mu) + \int_{\hat{\mu}_{uB}}^{\bar{\mu}} \int_{\underline{\theta}}^{\hat{\theta}_u} \mu \theta dF_u(\theta) dG(\mu).$$

Thus, a university maximizes the total wage bill of its student body by maximizing the total product of all of its students who are employed at good jobs.

An immediate implication is that some ‘‘A’’ students from each university always receive good jobs. A student only receives a good job if his expected productivity exceeds the endogenous level K (associated with measure Γ of students receiving good jobs). The common support assumptions on academic abilities and social skills ensures that both universities have some of the most able students. By setting a sufficiently high grading standard, a university can ensure that its ‘‘A’’ students have productivities arbitrarily close to $\bar{\theta}$, and some of these students will also have high social skills and hence receive good jobs.

3.2.1 Equilibrium

A symmetric pure strategy equilibrium is a collection of grading standards $(\hat{\theta}_H^*, \hat{\theta}_I^*)$, social skill cutoffs, $\hat{\mu}_{ug}^*$, $u \in \{H, I\}$, $g \in \{A, B\}$, and minimal productivity standard, K^* , such that:

- $\hat{\theta}_u^*$ maximizes the total productivity of students at a type $u \in \{H, I\}$ university who receive good jobs given productivity standard K^* ;
- Firms assign good jobs to maximize profit:

- $\underline{\mu}E_{ug}[\theta|\hat{\theta}_u^*] \geq K^*$, if $\mu_{ug}^* = \underline{\mu}$ and all ug students receive good jobs;
- $\bar{\mu}E_{ug}[\theta|\hat{\theta}_u^*] \leq K^*$, if $\mu_{ug}^* = \bar{\mu}$ and no ug students receive good jobs;
- $\hat{\mu}_{ug}^*E_{ug}[\theta|\hat{\theta}_u^*] = K^*$, otherwise.

- K^* “clears” the market: given $\hat{\theta}_H^*$, $\hat{\theta}_I^*$ and $\hat{\mu}_{ug}^*$, the measure of students with expected productivity of at least K^* is Γ :

$$\begin{aligned} & \alpha \left[(1 - G(\hat{\mu}_{HA}^*)) (1 - F_H(\hat{\theta}_H^*)) + (1 - G(\hat{\mu}_{HB}^*)) F_H(\hat{\theta}_H^*) \right] + \\ & + (1 - \alpha) \left[(1 - G(\hat{\mu}_{IA}^*)) (1 - F_I(\hat{\theta}_I^*)) + (1 - G(\hat{\mu}_{IB}^*)) F_I(\hat{\theta}_I^*) \right] = \Gamma. \end{aligned} \quad (3.1)$$

The next proposition establishes the existence of a symmetric equilibrium (i.e., each type u school sets the same grading standard). Optimization by schools further pins down the product of students at good jobs—each university will choose a grading standard that maximizes the productivity of students who receive good jobs.

Proposition 9 *A pure strategy symmetric equilibrium exists. There is a unique equilibrium productivity standard K for a good job, and the expected product of students from school type u who receive good jobs in equilibrium is unique.*

Uniqueness of the grading standard is not guaranteed absent assumptions on the measure of good jobs. For example, if there were so many good jobs that every student from a type H university receives one, then a type H university could achieve this either by giving all students “A”s, or by giving a very few students “A”s, so that the expected productivity of its “B” students was high enough that they receive good jobs.

3.3 No Heterogeneity in Social Skills

We begin by analyzing the special case in which all students have the same social skills, $\mu = 1$. To solve for equilibrium, first notice that it is never an equilibrium for some, but not all, students with grade g from university u to get good jobs. Were this so, a university could marginally increase its grading standard, raising the average ability of both its “A” and “B” students. In turn, the expected productivity of all students with grade g is raised, so that all now receive good jobs, increasing total productivity of u alumni at good jobs. Moreover, for an equilibrium to exist in which “B” students receive good jobs, Γ must exceed the measure α of students at H universities. Therefore, as long as $\Gamma \leq \alpha$, only “A” students receive good jobs in equilibrium when there is no variation in social skills.

So consider a *symmetric* equilibrium in which only “A” students receive good jobs. Were, say, HA students to have higher expected academic productivities than IA students, then a type H university could lower its standard for an “A” slightly, and more of its students would then receive good jobs. Therefore, in equilibrium, the expected productivity of “A” students at the two types of schools must be equal. Combined with the capacity constraint, this implies that equilibrium is fully described by:

$$\frac{\int_{\hat{\theta}_H^*}^{\bar{\theta}} \theta f_H(\theta) d\theta}{\int_{\hat{\theta}_H^*}^{\bar{\theta}} f_H(\theta) d\theta} = \frac{\int_{\hat{\theta}_I^*}^{\bar{\theta}} \theta f_I(\theta) d\theta}{\int_{\hat{\theta}_I^*}^{\bar{\theta}} f_I(\theta) d\theta} \quad (3.2)$$

$$\alpha \int_{\hat{\theta}_H^*}^{\bar{\theta}} f_H(\theta) d\theta + (1 - \alpha) \int_{\hat{\theta}_I^*}^{\bar{\theta}} f_I(\theta) d\theta = \Gamma. \quad (3.3)$$

The next result establishes that under a weak sufficient condition, type H universities set slacker grading standards than type I universities.

Proposition 10 *Suppose that Γ is small enough that in equilibrium not all H students receive good jobs. Then, in equilibrium, $\hat{\theta}_H^* < \hat{\theta}_I^*$.*

Proposition 10 implies that there is a positive mass of students at type H universities who receive “A”s, but whose ability is low enough that they would receive “B”s at a type I university. This asymmetry simply says that better universities have incentives to dilute the mass of good “A” students with students who would receive “B”s elsewhere: even after dilution, the better upper tail of good students at type H universities makes an average “A” student at a type H university as good as a typical “A” student from a type I university.

The result that $\hat{\theta}_H^* < \hat{\theta}_I^*$ holds *unless every* student at a type H university is assigned to a good job. In particular, $\Gamma \leq \alpha$ is far from necessary for this result, as in equilibrium a positive measure of students from type I universities receive “A” grades. This result compares grading standards in equilibrium of two types of universities, and does not depend on how many university types are there.

Social Planner. A social planner can set and enforce grading standards at schools (say, by limiting the quantity of “A” grades), but cannot assign particular students to particular jobs. She seeks to maximize total output in the economy, setting grading standards to solve

$$\max_{\hat{\theta}_H, \hat{\theta}_I} \alpha \int_{\hat{\theta}_H}^{\bar{\theta}} \theta dF_H(\theta) + (1 - \alpha) \int_{\hat{\theta}_I}^{\bar{\theta}} \theta dF_I(\theta) \quad (3.4)$$

$$\text{s.t. } \alpha \int_{\hat{\theta}_H}^{\bar{\theta}} dF_H(\theta) + (1 - \alpha) \int_{\hat{\theta}_I}^{\bar{\theta}} dF_I(\theta) = \Gamma. \quad (3.5)$$

Proposition 11 *The social planner sets a common grading standard, $\hat{\theta}_H^P = \hat{\theta}_I^P = \hat{\theta}^P$.*

When there is no heterogeneity in social skills, no inefficiencies are created by the fact that with the same grading standard, the expected productivity of “A” students is higher at a type H university: the “right” students receive good job assignments.

Corollary 2 *Suppose Γ is small enough that not all H students receive good jobs. Then $\hat{\theta}_H^* < \hat{\theta}^P < \hat{\theta}_I^*$: in equilibrium, too many H students and too few I students receive good jobs.*

The corollary follows since the good jobs capacity is exhausted in both the social optimum and equilibrium, so both $\hat{\theta}_H^* < \hat{\theta}_I^* \leq \hat{\theta}^P$ and $\hat{\theta}^P \leq \hat{\theta}_H^* < \hat{\theta}_I^*$ cannot occur.

Figure 3.1 conveys the intuition, depicting equilibrium and social planner outcomes. CFOSD implies that the solid line of equal expected productivities lies below the 45° line associated with equal grading standards; Proposition 10 says that the intersection of the dashed capacity constraint line and the solid equal productivity line is below the equal grading standards line; and the capacity constraint must be negatively sloped. Strategic considerations not only induce better universities to set slacker grading standards than is socially optimal, but they also force lesser universities to set stricter standards in order to compete. Still, type I universities prefer this outcome to one in which no grades are disclosed, in which case no I alumni would receive good jobs.

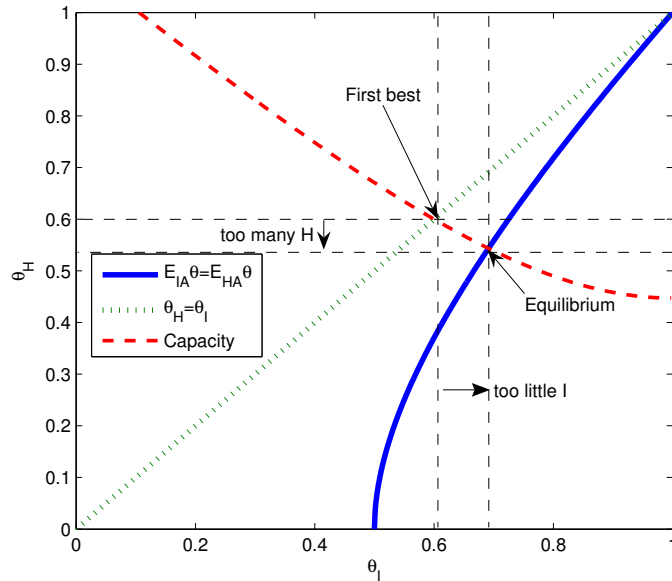


Figure 3.1: Comparison of Equilibrium and Social Planner outcomes.

We next characterize how the primitives of our economy affect outcomes.

More type H universities. An increase in the fraction α of top schools flattens the dashed capacity

constraint line because changes in grading standards at type H universities now have bigger impacts on the measure of students with A s. There are conflicting effects: increasing α shifts the composition of schools from those with high standards to those with low standards; but increasing α increases the supply of able students. To determine whether grading standards improve, observe that a marginal increase in α is a (not necessarily shape preserving) counterclockwise rotation of the capacity line, and one point of the line remains the same. Whenever this point is above the 45° line, the intersection with the (blue) equal productivity line shifts up and grading standards rise. Moreover, when $\alpha = \frac{1}{2}$, the rotation point is above 45° line, since $1 - F_H(t) > 1 - F_I(t)$. Thus, there exists a $\bar{\alpha} < \frac{1}{2}$ such that for $\alpha > \bar{\alpha}$, increasing the fraction of type H schools raises grading standards. That is, the supply effect dominates the composition effect when there are enough top universities.

Improvements in student body composition. Improving the distribution of student abilities at a type I university causes type H universities to set higher standards for an “A”, but has ambiguous effects on I ’s grading standard. To see this, notice that the “equal expected productivity” line shifts toward the 45° line (productivity of I university improves), and the capacity line moves to the right (for the same quantity of “A” students at a H university, one now needs a higher grading standard in I to fill the fewer remaining good jobs). As a result, the equilibrium $\hat{\theta}_H^*$ must rise, but the effect on $\hat{\theta}_I^*$ is ambiguous—the number of I students who receive good jobs must rise, but whether $\hat{\theta}_I^*$ increases depends on whether or not the increased quality composition of good I students dominates the effect of more good I students.

Intuitively, improving the distribution at type I schools creates added “competition” for type H schools forcing them to raise standards. The ambiguous impact on type I schools reflects that (a) type I schools can lower grading standards and still have a higher average quality of “A” students, but (b) the better distribution also increases competition for type I schools, raising the average ability required for a good job placement. Analogously, one can show that a deterioration in the distribution of abilities at type H universities eases grading standards at type I universities, but has ambiguous effects at type H schools.

Good Jobs, Grading Standards and Grade Inflation. We next characterize how increases in the number of good jobs affect equilibrium outcomes. It follows directly that increasing Γ causes all universities to lower grading standards: this reduction in the average quality of “A” students implies there is grade inflation. In particular, increasing Γ shifts the capacity line outward away from the top right corner, shifting equilibrium outcomes away along the “equal expectations about “A” students” locus line. The social planner’s choice shifts away along the 45° line, resulting in an equal decrease in grading standards and an increase in “A” grades at both universities.

We are especially interested in identifying when equilibrium grading standards fall by more at top schools, and when this translates into higher grade inflation at top schools. Define $Q_u(t) = E(\theta|u, \theta > t)$, where

$Q(\bar{\theta}) = \bar{\theta}$ preserves continuity. We characterize the relative impact of Γ on grading standards via the implicit function $\hat{\theta}_H(\hat{\theta}_I)$ defined by $Q_H(\hat{\theta}_H) = Q_I(\hat{\theta}_I) = K$, by varying K (K falls with Γ): grading standards fall faster at type H schools than type I schools if and only if $\hat{\theta}'_H(\hat{\theta}_I) > 1$.

Proposition 12 *Suppose there are sufficiently few good jobs, Γ . Then a slight increase in Γ causes grading standards to fall faster at type H schools than type I schools.*

One would like to extend this result to settings where the number of good jobs is larger, maintaining only the premise that Γ is not so large that all H students receive good jobs. To do this, we consider a family of ability densities with linear right tails, where the linear right tail is “long” enough that it describes the abilities of A students:

$$f_u(\theta|a_u, b_u) = a_u + b_u\theta, \theta \in [t, 1] \text{ for some } t, \quad (3.6)$$

where, to ease presentation, we assume a $[0, 1]$ support (extensions to a support $[\underline{\theta}, \bar{\theta}]$ are routine). Positivity of the density implies that $a_u + b_u > 0$ and $a_u + b_u t > 0$, and $f_H(\cdot) \succeq_C f_I(\cdot)$ implies that $a_H b_I < a_I b_H$. We also need that b_H sufficiently exceeds b_I . When densities are linear on their full support, $b_H > b_I$ suffices.

Proposition 13 *Consider densities with linear right tails, where b_H sufficiently exceeds b_I , so that $a_H b_I(1 + \hat{\theta}_I) < a_I b_H(1 + \hat{\theta}_H)$. Then if $\hat{\theta}_u > t$, an increase in Γ causes grading standards to fall faster at type H universities than type I universities.*

Corollary 3 *If $f_H(\hat{\theta}_H) \geq f_I(\hat{\theta}_I)$ and grading standards fall faster in H , then the number of “A”s increases faster at type H universities than type I universities.*

For example, if $f_H(t)/f_I(t)$ increases in t , then there exists a \tilde{t} such that $f_H(t') > f_I(t'), \forall t' > \tilde{t}$, in which case Corollary 3 follows if and only if there are sufficiently few good jobs, Γ .

In sum, when students only differ in academic ability, universities with better student bodies press some students from lesser universities out of good job assignments by setting slacker grading standards than is socially optimal. Further, under plausible scenarios, more good jobs causes greater grade inflation at top universities. We now explore how heterogeneity in social skills affects these conclusions.

3.4 Heterogeneous Social Skills

Suppose now that students differ in their social skills. For heterogeneity in social skills to alter outcomes, there must be a sufficient difference between the highest and lowest social skill that not all “A” students receive good job assignments. If not, then our previous analysis characterizes outcomes. We maintain the

assumption that the dispersion is not sufficient for “B” students to receive good job assignments in equilibrium. Thus, the support $[\underline{\mu}, \bar{\mu}]$ of the distribution of social skills $G(\cdot)$ is neither very small, nor very large; i.e., heterogeneity in social skills is “intermediate” so that $K^* > \max\{\underline{\mu}E_u[\theta|\theta > \hat{\theta}_u^*], \bar{\mu}E[\theta|\theta < \hat{\theta}_u^*]\}$.

Proposition 14 *Suppose $\mu g(\mu)$ is increasing in μ and heterogeneity in social skills is intermediate. Then, in the unique equilibrium, type I schools set higher grading standards than type H schools. Further, $\hat{\mu}_H^* \leq \hat{\mu}_I^*$, where the inequality is strict as long as some A students do not receive good jobs.*

The result says that as long as the density over social skills does not fall quickly, $g'(\mu) > -\frac{g(\mu)}{\mu}$, then top universities set slacker grading standards. When Γ is small enough that in equilibrium not all students at a school receive “A”s, and the dispersion in social skills is sufficient that an “A” student with the lowest social skill is not offered a good job, then equilibrium is characterized by interior solutions, and a university’s best response is characterized by the intersection of

$$\hat{\theta}_u = K \frac{\frac{g(\hat{\mu}_u)}{1-G(\hat{\mu}_u)}\hat{\mu}_u}{E[\mu|\mu > \hat{\mu}_u] + \frac{g(\hat{\mu}_u)}{1-G(\hat{\mu}_u)}\hat{\mu}_u^2} \equiv R(\hat{\mu}_u)$$

$$\hat{\mu}_u E_u[\theta|\theta > \hat{\theta}_u] = K \Leftrightarrow \hat{\mu}_u Q(\hat{\theta}_u) = K.$$

The bottom efficient job assignment equation is decreasing in $(\hat{\mu}, \hat{\theta})$ space (see Figure 3.2). The CFOSD assumption implies that $E_I[\theta|\theta > \hat{\theta}] < E_H[\theta|\theta > \hat{\theta}]$, for every $\hat{\theta}$. Therefore, $\hat{\mu}_{HA}^* < \hat{\mu}_{IA}^*$. Since $\mu g(\mu)$ is increasing in μ , the right-hand side of the top equation is increasing in μ , which combined with $\hat{\mu}_{HA}^* < \hat{\mu}_{IA}^*$, implies $\hat{\theta}_H^* < \hat{\theta}_I^*$.

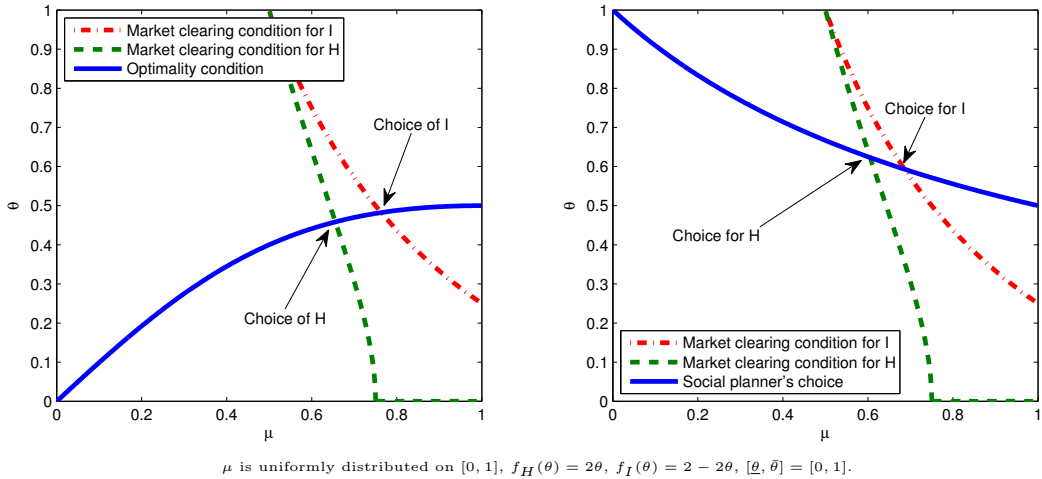


Figure 3.2: Equilibrium and Social Planner outcomes.

Corollary 4 *In any interior equilibrium, $E[\theta|H, \theta > \hat{\theta}_H^*] > E[\theta|I, \theta > \hat{\theta}_I^*]$.*

The corollary follows directly from optimal job assignment, $\hat{\mu}_{HA}^* E[\theta|H, \theta > \hat{\theta}_H^*] = \hat{\mu}_{IA}^* E[\theta|I, \theta > \hat{\theta}_I^*]$; and Proposition 14, which states that $\hat{\mu}_{HA}^* < \hat{\mu}_{IA}^*$. Corollary 4 says that otherwise identical students from H and from I receive different wages at good jobs: in particular, students from lesser universities receive lower wages.

Indeed, it also follows that the *average* wage of a student of better university employed on a good job is also higher:

$$\begin{aligned}
& E[\theta\mu|\text{employed from H}] - E[\theta\mu|\text{employed from I}] \\
&= E[\mu|\mu > \hat{\mu}_H^*]E_H[\theta|\theta > \hat{\theta}_H^*] - E[\mu|\mu > \hat{\mu}_I^*]E_I[\theta|\theta > \hat{\theta}_I^*] \\
&= K \left(\frac{E[\mu|\mu > \hat{\mu}_H^*]}{\hat{\mu}_H^*} - \frac{E[\mu|\mu > \hat{\mu}_I^*]}{\hat{\mu}_I^*} \right) \\
&= \frac{K}{\hat{\mu}_H^* \hat{\mu}_I^*} (E[\mu|\mu > \hat{\mu}_H^*](\mu_I^* - \mu_H^*) - \mu_H^* (E[\mu|\mu > \hat{\mu}_I^*] - E[\mu|\mu > \hat{\mu}_H^*])) > 0.
\end{aligned}$$

The last inequality follows since $g(\mu)$ features an increasing hazard, so $E[\mu|\mu > \hat{\mu}_I^*] - E[\mu|\mu > \hat{\mu}_H^*] < \hat{\mu}_I^* - \hat{\mu}_H^*$.

We now derive the qualitative impact of intermediate levels of heterogeneity in social skills on the social planner's grading standards.

Proposition 15 *Suppose heterogeneity in student social skills is intermediate, so that not all “A” students receive good jobs, and no “B” students receive good jobs. Then the social planner sets $\hat{\theta}_H^P > \hat{\theta}_I^P$ and $\hat{\mu}_{HA}^P < \hat{\mu}_{IA}^P$.*

The social planner sets $\hat{\mu}_{HA}^P E[\theta|\theta > \hat{\theta}_H^P] = \hat{\mu}_{IA}^P E[\theta|\theta > \hat{\theta}_I^P]$ and $E[\mu|\mu > \hat{\mu}_{HA}^P] \hat{\theta}_H^P = E[\mu|\mu > \hat{\mu}_{IA}^P] \hat{\theta}_I^P$. The first equality confirms that a social planner sets the same standards on the social skills of students assigned to good jobs as competitive firms: each marginal student has the same expected productivity given his social skills and the information contained in an “A” grade from his university. As a result, even though the distributions of social skills at universities are the same, the average social skill of “A” students from a type H university who receive good job assignments is less. The second equality says that the social planner sets grading standards to equate the expected productivities of the marginal “A” student at each university. Then, because the average social skill of students from a type H university with good jobs is less, a social planner sets a more demanding grading standard at type H universities.

Heterogeneity across universities in student body compositions leads to worse equilibrium outcomes than were there a common ability distribution at schools. One might conjecture that this heterogeneity also hinders the social planner because it causes firms to distort hiring decisions toward students from better universities. This conjecture is false: homogeneity harms a planner's ability to distinguish better populations of students. To see this, note that with heterogeneous distributions, the social planner *could* set

common grading and social skill cutoffs; however, the planner chooses not to. It follows that homogenizing the university pool is suboptimal.

In sum, as long as there is not so much heterogeneity in social skills that some “B” students receive good jobs, then under mild conditions, better universities set lower grading standards, even though a social planner would make the opposite choice. These findings extend when some “B” students with exceptional skills also receive good jobs, as long as the distribution of social skills is such that “A” students dominate decision making. We believe that this is the relevant real world scenario. However, if, for example, $g(\mu)$ is sufficiently flat with sufficient dispersion, then type H schools may weigh the job prospects of “B” students by enough that they set higher grading standards, in order to raise the fraction of “B” students that receive good jobs.

3.5 Conclusion

The central message of this paper is that competition for good job assignments for graduates causes better universities to set lower standards for “A”s, because their marginal “A” students can ride on the coat tails of the better average qualities of “A” students. We show that a social planner sets the opposite ordering on grading standards. We also show that increases in the number of good jobs drives down standards for “A”s, and that under plausible scenarios, standards fall more at better schools.

Although our setting features just two types of universities, *none* of our analysis hinges on this modeling choice. To see that our results extend as long as university types can be CFOSD-ordered by their distributions over academic abilities, note that additional types enter via the equilibrium standard for a good job assignment, and our analysis simply establishes that better schools set lower grading standards. In particular, our analysis extends when there is a continuum of different university types. So, too, our analysis does not hinge on our common support assumption, as long as some students at a lesser schools get good jobs.

Another feature of our model is two grades. Increasing the quantity of grades discretely is cumbersome but not too fruitful: our model immediately extends to the case of three grades if the relevant competition on the labor market is between “A” students of different university types. To address concerns over robustness of our model to continuous grades, consider a model where students’ GPA is an average of binary “pass-fail” grades given for multiple courses. Professors face idiosyncratic noise when evaluating the students’ ability endowment generated by imperfections in professors’ assessment of students’ ability, matching between student’s and professor’s teaching style, random exam performance and other factors. University mandates how many good grades professors can award to students. Then, as precision of professor’s assesment ability increases, the outcome will be arbitrarily close to the outcome of our two-grade model. Even in the setting

where noise is positive, firms, knowing how much good grades universities suggest to award, will hire for good jobs top GPA students until the expected productivity of hired students from university of each type is equal to the market-suggested. As a result, even with noise, firms will hire more students from better universities than it is socially optimal, because better schools will allow to award more good grades than is socially optimal. We decided to forego this complication as it does not help the understanding of the economics behind the exploitation of better student body by better universities.

One feature that we do not integrate into our model is effort, and how grading standards affect student effort, where effort both affects academic performance and on-the-job productivity. In such a setting, universities must account for how their grading standards affect effort choices and equilibrium job assignment. Lower standards for “A” grades may induce students to exert less effort (especially if the density of ability for the marginal student is low), and universities will internalize this effect. One can clearly provide conditions under which endogenizing effort does not reverse the conclusion that better universities set lower grading standards. A thorough analysis of effort and grading interaction in equilibrium is an interesting topic for future research.

Appendix A

Fraternities and Labor Market Outcomes

Proof of Proposition 3: First observe that in light of Proposition 2, if (θ, μ) applies to the fraternity in equilibrium, then so does (θ, μ') . Suppose the proposition were false. Then for $\varepsilon > 0$, sufficiently small, the set

$$\Theta_\varepsilon = \{\theta | \exists z_\theta : \int_{-\infty}^{z_\theta} a(\theta, x)b(\theta, x)h_\mu(x)dx > \varepsilon, \int_{z_\theta}^{+\infty} (1 - b(\theta, x)) h_\mu(x)dx > \varepsilon\}$$

has positive measure, i.e., there exists $\delta > 0$ such that $\int_{\Theta_\varepsilon} h_\theta(x)dx > \delta$. For every θ in Θ_ε pick a set of $K = \{(\theta, \mu) | \mu < z_\theta, a(\mu, \theta)b(\mu, \theta) = 1\}$ and $L = \{(\theta, \mu) | \mu > z_\theta, b(\mu, \theta) = 0\}$ such that $\int_K a(\theta, \mu)b(\theta, \mu)h_\mu(x)dx = \int_L a(\theta, \mu)(1 - b(\theta, \mu))h_\mu(x)dx = \varepsilon$. But then the fraternity decision rule

$$\hat{b}(\theta, \mu) = b(\theta, \mu)(1 - I((\theta, \mu) \in K)) + I((\theta, \mu) \in L)$$

strictly raises the fraternity's payoff as $E(\mu|K) < E(\mu|L)$ and the expected wages generated by members and fraternity size are unchanged. Therefore, b could not have been an equilibrium strategy for the fraternity. \square

Proof of Lemma 1: Consider two signals, $x > y \in \tilde{\Theta}$, and two productivities, $\theta_2 > \theta_1 \in \Theta$, such that (x, θ_1) , (x, θ_2) and $(y, \theta_2) \in D$. By the MLRP property,

$$\frac{f(\theta_2|x)}{f(\theta_1|x)} > \frac{f(\theta_2|y)}{f(\theta_1|y)}.$$

Notice that for every $(j, k) \in D$, $f(j|k, D) = \frac{f(j|k)I((j,k) \in D)}{\int_{\Theta} I((\theta, \theta) \in D)dF(\theta|\theta)} = f(j|k) \frac{I((j,k) \in D)}{P(D)}$. Rewrite the MLRP condition:

$$\frac{f(\theta_2|x, D)}{f(\theta_1|x, D)} = \frac{\frac{f(\theta_2|x)I((\theta_2,x) \in D)}{P(D)}}{\frac{f(\theta_1|x)I((\theta_1,x) \in D)}{P(D)}} = \frac{\frac{f(\theta_2|x)1}{P(D)}}{\frac{f(\theta_1|x)1}{P(D)}} > \frac{\frac{f(\theta_2|y)1}{P(D)}}{\frac{f(\theta_1|y)1}{P(D)}} \geq \frac{\frac{f(\theta_2|y)I((\theta_2,y) \in D)}{P(D)}}{\frac{f(\theta_1|y)I((\theta_1,y) \in D)}{P(D)}} = \frac{f(\theta_2|y, D)}{f(\theta_1|y, D)}.$$

Therefore, if the MLRP condition holds for the entire support, it holds for a subset D of that support. This

condition ensures that $E(\theta|\tilde{\theta}, D)$ is an increasing function of $\tilde{\theta}$. By $F(x|\theta_2, D) \succeq F(x|\theta_1, D)$,

$$\begin{aligned} E_{\tilde{\theta}}[E(\theta|\tilde{\theta}, D)|\theta_2, D] &= \int_{\tilde{\Theta}} E(\theta|\tilde{\theta}, D)dF(\tilde{\theta}|\theta_2, D) \\ &> \int_{\tilde{\Theta}} E(\theta|\tilde{\theta}, D)dF(\tilde{\theta}|\theta_1, D) = E_{\tilde{\theta}}[E(\theta|\tilde{\theta}, D)|\theta_1, D]. \quad \square \end{aligned}$$

Proof of Proposition 5: We first show that if $n\bar{\mu} - c \geq \underline{w} - \underline{\theta}$, then the empty fraternity cannot be an equilibrium. If no one joins the fraternity, then equilibrium demands that $\underline{\theta}$ expect wage \underline{w} if he does not join; and the expected wages of students with ability greater than $\underline{\theta}$ who do not join exceed \underline{w} . With the full support assumption, following any signal realization $\tilde{\theta}$, firms can hold equilibrium beliefs that the anyone who joins the fraternity and generated that signal has ability $\underline{\theta}$. These beliefs minimize the wage of any student who joins the fraternity. Given these beliefs, since expected wages are continuous in θ , if $n\bar{\mu} - c + \underline{\theta} > \underline{w}$, then all students in a sufficiently small neighborhood of $(\underline{\theta}, \bar{\mu})$ would apply for fraternity membership, and since their measure is less than Γ , the fraternity would accept them. Hence, the empty fraternity cannot be an equilibrium.

Conversely, if $n\bar{\mu} - c \leq \underline{w} - \underline{\theta}$, then given the pessimistic beliefs by firms, $w_C(\tilde{\theta}) = \underline{\theta}$ so that $(\underline{\theta}, \bar{\mu})$ at least weakly prefers not to apply to the fraternity; and all other types strictly prefer not to apply. Hence, no one applying to the fraternity is an equilibrium. \square

Proof of Proposition 6: To prove existence, it suffices to characterize student and fraternity choices via the continuous functions $\mu_A(\cdot)$ and $\mu_B(\cdot)$ (see Propositions 2 and 4), proving the existence of an equilibrium in which a student type (θ, μ) is a member of the fraternity if and only if $\mu \geq \max\{\mu_A(\theta), \mu_B(\theta)\}$. In particular, given $w_C(\cdot)$ and $w_{\bar{C}}(\cdot)$, $\mu_A(\theta)$ solves equation (1.2) at equality, for $\mu_A(\theta) \in (\underline{\mu}, \bar{\mu})$. Since $\mu_A(\cdot)$ is uniquely defined, it follows that $\mu_B(\cdot)$ is uniquely defined. We have established that $\mu_j : [\underline{\theta}, \bar{\theta}] \rightarrow [\underline{\mu}, \bar{\mu}]$, $j = A, B$, is continuous. The space of such functions, endowed with the weak* topology, is compact. So, too, we can focus on beliefs by firms about which student types are fraternity members that are summarized by continuous functions $\hat{\mu}_A(\cdot)$ and $\hat{\mu}_B(\cdot)$ about which student types apply and which ones are accepted by the fraternity, where (θ, μ) is a conjectured fraternity member if and only if $\mu \geq \max\{\hat{\mu}_A(\theta), \hat{\mu}_B(\theta)\}$. These beliefs, $\hat{\mu}_A(\cdot), \hat{\mu}_B(\cdot)$ then determine competitive wage functions,

$$(\hat{w}_C(\tilde{\theta}), \hat{w}_{\bar{C}}(\tilde{\theta})) = (E[\theta|\tilde{\theta}, \mu \geq \max\{\hat{\mu}_A(\theta), \hat{\mu}_B(\theta)\}], E[\theta|\tilde{\theta}, \mu < \max\{\hat{\mu}_A(\theta), \hat{\mu}_B(\theta)\}]),$$

and these wage functions, in turn, imply optimal student and fraternity best response choices, $\mu_A(\cdot), \mu_B(\cdot)$. Hence, we have a mapping from $(\hat{\mu}_A(\theta), \hat{\mu}_B(\theta))$ to $(\mu_A(\theta), \mu_B(\theta))$. Equilibrium is given by a fixed point to this mapping from [conjectured by firms] optimal student and fraternity choices to the best response optimal

student and fraternity choices; and we have just established that this mapping satisfies the conditions of Kakutani's fixed point theorem. \square

Proof of Lemma 2: $\mu_A(\theta)$ solves

$$E(w_C(\tilde{\theta})|\theta) + n\mu_A(\theta) - c = E(w_{\bar{C}}(\tilde{\theta})|\theta) \Leftrightarrow E(w_C(\tilde{\theta})|\theta) - E(w_{\bar{C}}(\tilde{\theta})|\theta) = c - n\mu_A(\theta).$$

Then $\mu_A(\theta)$ is increasing in θ if and only if

$$\frac{\partial}{\partial \theta} E(w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta})|\theta) = \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta}) f_{\tilde{\theta}}(\tilde{\theta}|\theta) d\tilde{\theta} < 0.$$

Consider $g(\tilde{\theta}|\theta) \equiv \frac{\partial}{\partial \theta} f_{\tilde{\theta}}(\tilde{\theta}|\theta)$. As $g(\tilde{\theta}|\theta)$ is the change in the distribution of signal $\tilde{\theta}$ due to a change in θ , the integral of g over the support of $\tilde{\theta}$ is zero. The MLRP assumption implies that $g(\tilde{\theta}|\theta)$ is increasing in $\tilde{\theta}$:

$$\text{MLRP: } \frac{f_{\tilde{\theta}}(t_2|\theta + \Delta_\theta)}{f_{\tilde{\theta}}(t_2|\theta)} > \frac{f_{\tilde{\theta}}(t_1|\theta + \Delta_\theta)}{f_{\tilde{\theta}}(t_1|\theta)} \quad \forall t_2 > t_1, \Delta_\theta > 0 \text{ implies}$$

$$\frac{1}{f_{\tilde{\theta}}(t_2|\theta)} \frac{\Delta f_{\tilde{\theta}}(t_2|\theta)}{\Delta_\theta} > \frac{1}{f_{\tilde{\theta}}(t_1|\theta)} \frac{\Delta f_{\tilde{\theta}}(t_1|\theta)}{\Delta_\theta}$$

Therefore,

$$\frac{\partial}{\partial \theta} \ln f_{\tilde{\theta}}(t_2|\theta) \geq \frac{\partial}{\partial \theta} \ln f_{\tilde{\theta}}(t_1|\theta)$$

implying that $g(\tilde{\theta}|\theta)$ is nondecreasing in $\tilde{\theta}$; and since g integrates to zero, there exists a K such that $g(\tilde{\theta}|\theta) \leq 0$ for $\tilde{\theta} \leq K$, and $g(\tilde{\theta}|\theta) \geq 0$ for $\tilde{\theta} > K$.

Remember, we want to establish when

$$\frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} [w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta})] f_{\tilde{\theta}}(\tilde{\theta}|\theta) d\tilde{\theta} = \int_{-\infty}^{\infty} [w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta})] \frac{\partial f_{\tilde{\theta}}(\tilde{\theta}|\theta)}{\partial \theta} d\tilde{\theta} < 0$$

Subtracting $[w_C(K) - w_{\bar{C}}(K)] \int_{-\infty}^{\infty} \frac{\partial f_{\tilde{\theta}}(\tilde{\theta}|\theta)}{\partial \theta} d\tilde{\theta} = 0$ from the integral and breaking the multiplicands under the integral into two parts yields

$$\begin{aligned} & \int_{-\infty}^{\infty} \left((w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta})) - (w_C(K) - w_{\bar{C}}(K)) \right) \frac{\partial f_{\tilde{\theta}}(\tilde{\theta}|\theta)}{\partial \theta} d\tilde{\theta} \\ &= \int_{-\infty}^K \left((w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta})) - (w_C(K) - w_{\bar{C}}(K)) \right) g(\tilde{\theta}|\theta) d\tilde{\theta} \end{aligned}$$

$$+ \int_K^\infty \left((w_C(\tilde{\theta}) - w_{\bar{C}}(\tilde{\theta})) - (w_C(K) - w_{\bar{C}}(K)) \right) g(\tilde{\theta}|\theta) d\tilde{\theta}.$$

From the premise that $w_C(x) - w_{\bar{C}}(x)$ is decreasing, the difference in the w terms in the first integral is positive, and negative in the second; by construction, the g term in the first integral is negative, and positive in the second integral. Therefore, the integral is the sum of two negative values.

Finally, inspection reveals that an analogous argument holds if there are a finite number of signals, $\tilde{\theta}$, interpreting the signal density $f_{\tilde{\theta}}(\tilde{\theta}|\theta)$ as the probability mass on signal $\tilde{\theta}$. \square

Proof of Proposition 7: Denote the set of students with $\theta < \theta_0$ as P , the set of students in fraternity as P_2 , and the rest P_1 ; and let $m(\cdot)$ be the measure of students in the argument set. Then

$$w_C(\tilde{\theta}) = E(\theta|\tilde{\theta}, P_2) \quad \text{and} \quad w_{\bar{C}}(\tilde{\theta}) = E(\theta|\tilde{\theta}, P \cup P_1).$$

Rewrite $w_{\bar{C}}$ as

$$w_{\bar{C}}(\tilde{\theta}) = \frac{m(P)}{m(P) + m(P_1)} E(\theta|\tilde{\theta}, P) + \frac{m(P_1)}{m(P) + m(P_1)} E(\theta|\tilde{\theta}, P_1).$$

Take two signals, H and L , with $H > L$. Then the expected wage premium is decreasing if

$$\begin{aligned} & E(\theta|H, P_2) - \left[\frac{m(P)}{m(P) + m(P_1)} E(\theta|H, P) + \frac{m(P_1)}{m(P) + m(P_1)} E(\theta|H, P_1) \right] \\ < & E(\theta|L, P_2) - \left[\frac{m(P)}{m(P) + m(P_1)} E(\theta|L, P) + \frac{m(P_1)}{m(P) + m(P_1)} E(\theta|L, P_1) \right]. \end{aligned} \quad (\text{A.1})$$

Observe that $\theta_0 \rightarrow \bar{\theta}$ implies:

$$\begin{aligned} E(\theta|H, P_2) &\rightarrow \bar{\theta} \quad \text{and} \quad E(\theta|L, P_2) \rightarrow \bar{\theta} \\ \frac{m(P)}{m(P) + m(P_1)} E(\theta|H, P) + \frac{m(P_1)}{m(P) + m(P_1)} E(\theta|H, P_1) &\rightarrow E(\theta|H) \\ \frac{m(P)}{m(P) + m(P_1)} E(\theta|L, P) + \frac{m(P_1)}{m(P) + m(P_1)} E(\theta|L, P_1) &\rightarrow E(\theta|L). \end{aligned}$$

Therefore, as $\theta_0 \rightarrow \bar{\theta}$, (A.1) approaches

$$\bar{\theta} - E(\theta|H) < \bar{\theta} - E(\theta|L),$$

which holds as by the MLRP assumption, $E(\theta|H) > E(\theta|L)$. As the distribution of θ has full support, and there are no atoms in the distribution, $E(\theta|H, \theta > \theta_0)$ and $E(\theta|H, \theta < \theta_0)$ are continuous in θ_0 . Therefore, there exists a $\hat{\theta}_0(H, L) < \bar{\theta}$ such that for all $\theta_0 \geq \hat{\theta}_0(H, L)$, the expected wage premium of fraternity

members is decreasing in $\tilde{\theta}$. This bound on θ_0 depends on the signals H and L ; however, Assumption 1 ensures the existence of a uniform bound. In particular, if the support $\tilde{\theta}$ is finite, then the uniform bound is the maximum of the bounds for each signal pair; and if the support of $\tilde{\theta}$ is not finite, but the support of $f_{\tilde{\theta}}(\tilde{\theta}|\bar{\theta})$ is non-trivial, then the expected wage difference that $\bar{\theta}$ expects places strictly positive probability on signals bounded away from $\hat{\theta}$. Hence, there exists a small enough $\underline{\Gamma} > 0$ such that for $\Gamma < \underline{\Gamma}$, the equilibrium expected wage premium from fraternity membership is declining in ability. \square

Appendix B

Equilibrium Sovereign Default with Endogenous Exchange Rate Depreciation

B.1 Centralized Equilibrium

At the beginning of the game consumer chooses whether he wants to default or not is **Household's New Problem**:

$$U(b, y) = \max_{h \in \{0,1\}} hW(y) + (1-h)V(b, y).$$

The borrowers's problem conditional on not defaulting this period is **New Problem With No Default**:

$$\begin{aligned} V(b, y) &= \max_{c, x, m, b'} u(c, m) + \beta EU(b', y'), & (B.1) \\ \text{s.t.} & \\ c + x + b &= y + q(y, b')b', \\ m &= f(x), \\ \ln y' &= \rho \ln y + \epsilon, \quad \epsilon \sim \mathcal{N}(0, s^2). \end{aligned}$$

If country chooses to default, then its value function is a solution to **New Problem In Default**:

$$\begin{aligned} W(y) &= \max_{c, x, m} u(c, m) + \beta E(\phi W(y') + (1-\phi)U(0, y')), & (B.2) \\ \text{s.t.} & \\ c + x &= y + \Pi, \\ m &= (1-\pi)f(x), \\ \ln y' &= \rho \ln y + \epsilon, \quad \epsilon \sim \mathcal{N}(0, s^2). \end{aligned}$$

Combined with the Zero Profit Condition, solution to this problem will give the same values as the Equilibrium we want to study. However, it does not give the value of the exchange rate. We recover this from the first-order conditions to the New household choice problems:

$$e_V(y, b) = \frac{\partial u(c_V(y, b), m_V(y, b))}{\partial m} / \frac{\partial u(c_V(y, b), m_V(y, b))}{\partial c}$$

$$e_W(y) = \frac{\partial u(c_W(y), m_W(y))}{\partial m} / \frac{\partial u(c_W(y), m_W(y))}{\partial c}.$$

B.2 Estimation

To test our model's implications, we estimated various parameters using Argentine data. In many cases, we could have better characterized the Argentine data generating process with different parametric functional forms, but this would distract from the model we introduce. Our goal is not to match Argentine perfectly, though there is a certain virtue in doing so, but we fear that additional complexity will impede one's intuition for the model. Rather we decided to stay in the simple world of AR(1) processes and (relatively) linear functions to provide a reasonable test of the quantitative implications of our model .

Output Time Series

Output y was calibrated to quarterly deseasoned per-capita GDP in constant prices, which was assumed to follow AR(1) process. It was not detrended because trend seems to be too small. Regression equation is

$$\ln y_t - m = \rho(\ln y_{t-1} - m) + \epsilon_t, \epsilon_t \sim N(0, s^2).$$

m is the normalizing coefficient. Estimates are following:

Variable	Coefficient	(Std. Err.)
m	9.0669*	(0.1464)
ρ	0.9878*	(0.0081)
s	0.0258	

Table B.1: Estimation results, output dynamics, One star denotes 1% significance.

A Dickey-Fuller test does not reject a unit root hypothesis for this process, therefore, given standard deviations may be biased downward. Controlling a linear trend does not help to reject the unit root hypothesis, nor does it give much different estimates of other parameters.

B.2.1 Consumer's Utility Parameters

Utility function parameters were estimated from the first-order condition of consumer. c was taken to be equal to consumption (both private and public) plus investment from INDEC data. Nominal imports quantity was deduced from import value from INDEC data divided by the exchange rate, obtained from European Bank.

$$\ln e = \ln \frac{1 - \alpha}{\alpha} + (\kappa - 1) \left(\ln \frac{c/y}{m/y} \right).$$

Estimates are following:

Table B.2: Estimation results, output dynamics.

Variable	Coefficient	(Std. Err.)
$1 - \kappa$	0.1553*	(0.0440)
$\ln \frac{1-\alpha}{\alpha}$	-0.3472*	(0.0972)

Within a 95% confidence interval, we can conclude that $\kappa < 1$, so goods are not perfect substitutes, and $\alpha > 0.5$. With the obvious transformations of the estimation in Table B.2, $\kappa = 0.8447$ and $\alpha = 0.5859$. Again, the regression is not perfect. An ARIMA(2,1,1) model specification seems to perform better to correct for the nonstationarity for which exchange rates are notorious.

B.2.2 Import-Export Equation

The most interesting regression seems to be the trade equation. It was estimated by nonlinear least squares:

$$\ln m_t = \ln(1 - \pi)I(\text{punished at } t) + \ln \theta_1 + \theta \ln(x_t - \theta_0) + \epsilon_t.$$

Here we decided to allow the import-export equation to be nonlinear (θ not necessarily equal to 1), have a fixed cost (θ_0 is not necessarily equal to 0). We do not explore why it happens that firms can earn positive profits in equilibrium; we just allow estimates to signal us about that. These degrees of freedom are not necessary, and they don't drive our main result. However, they certainly help to achieve better fit of default responses.

For numbers on imports and exports, we smoothed quarterly fluctuations on the INDEC data using an HP filter Hodrick and Prescott (1997) with smoothing parameter 400.

The estimate of $\ln(1 - \pi)$ suggests that π is equal to 0.5 with surprising precision. We will use this number in model outcome calculation; to make sure we don't fall victim of regression's non-robustness, we also solve for the model with $\pi = 0.8$ and $\pi = 0.2$, so whatever is the real value of π , the real outcome will

Table B.3: Estimation results, import-export conditions.

Variable	Coefficient	(Std. Err.)
θ	0.2082*	(0.0763)
θ_0	0.0467*	(0.0070)
θ_1	0.1959*	(0.0409)
$\ln(1 - \pi)$	-0.6939*	(0.0256)

be lying in between these two models' outcomes. As for values of θ s, they are not too far away from the ones a person can get by estimating the 1993-2000 subsample of "no default."

We estimate this equation on a subsample of 1993-2003. The reason why we do not continue on the sample of 2004 and further is that it seems that Argentine does not have the same instant recovery of terms of trade after default as we have assumed in the model. Consider the time series of $u_t = \frac{\Delta m_t}{\Delta x_t} \frac{x_t}{m_t}$, a measure of elasticity of change in import with the change of export, presented on Figure B.1. One can see an approximately constant elasticity during most of 1990s, then a decline, a series of definite changes in structure of equation, and an increase in 2004, with a stable more than 2 elasticity after default. The big jump down in 2001 is what we try to capture with $(1 - \pi)$ multiplier. In this estimation, we allow for a period of adjustment, a process of re-establishing of connections lost in 2001. Imports do not only increase due to an increase in x , but also from a renewed efficiency of trading. Summarizing, we don't use the data after default because we believe that after forgiving the default import depends not only on export on the same period, but also on export on previous periods.

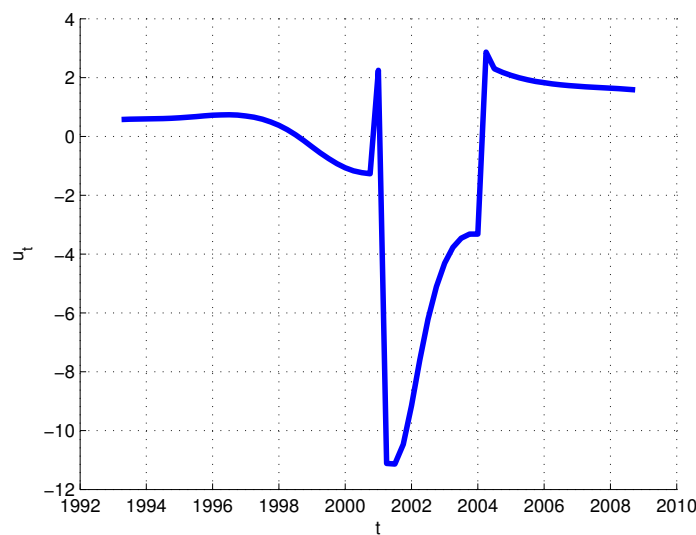


Figure B.1: Time series of $u_t = \frac{\Delta m_t}{\Delta x_t} \frac{x_t}{m_t}$, "sample elasticity."

B.3 Solving for First Best

Solving the “first best” general equilibrium is equivalent to solving the following convex programming problem, instead of the stochastic problem established above. Values for endogenous variables carry the subscript z_{fb}

$$V(\bar{y}) = \max_{c_{fb}, m_{fb}} \{ \alpha c_{fb}^\kappa + (1 - \alpha) m_{fb}^\kappa \}^{(1-\sigma)/\kappa} + \beta V(\bar{y}),$$

subject to:

$$c_{fb} + e_{fb} m_{fb} + b = \bar{y} + \Pi,$$

$$m_{fb} = \theta_1 (x_{fb} + \theta_0)^\theta.$$

The solution to which gives us

$$\bar{y} + \Pi = c_{fb} + e_{fb} \theta_1 (x_{fb} + \theta_0)^\theta, \quad (\text{B.3})$$

$$\alpha c_{fb}^{\kappa-1} = e_{fb}^{-1} (1 - \alpha) (x_{fb} + \theta_0)^{\theta(\kappa-1)}, \quad (\text{B.4})$$

$$m_{fb} = \theta_1 (x_{fb} + \theta_0)^\theta. \quad (\text{B.5})$$

The importer’s problem, gives two more conditions for the optimal quantity of exports and the level of profits.

$$\Pi_{fb} = e_{fb} m_{fb} - x_{fb},$$

$$m_{fb} = \theta_1 (x_{fb} + \theta_0)^\theta,$$

which gives optimal levels of x_{fb} , Π_{fb} :

$$x = (e_{fb} \theta_1 \theta)^{1/(\theta-1)} - \theta_0, \quad (\text{B.6})$$

$$\Pi = (e \theta_1)^{-1/(\theta-1)} (\theta^{-\theta/(\theta-1)} - \theta^{-1/(\theta-1)}) + \theta_0. \quad (\text{B.7})$$

Appendix C

University Competition, Grading Standards and Grade Inflation

Proof of Proposition 9. Substitute the market-clearing conditions into the maximized objectives of the universities. Each university maximizes a continuous function on the convex set $[\underline{\theta}, \bar{\theta}]$. Therefore, by Berge’s maximum theorem, there is an upper hemicontinuous best-response relation $\theta_u^*(K)$. Labor-market clearing implies that $K^*(\theta_H^*, \theta_I^*)$ is a continuous function of its arguments, since the densities are positive on their support. Substituting the best responses of the two types of universities into the market-clearing condition yields a upper hemicontinuous correspondence $K^*(K)$, defined on $[\underline{\theta}, \bar{\theta}]$ to itself, which has a fixed point by Kakutani’s theorem.

Now suppose there were multiple equilibrium standards, K_1, K_2 , for good job assignments, with $K_1 > K_2$. Then, choosing the same grading standard facing K_2 , each university can place more students at good jobs. But then the labor markets cannot clear for both K_1 and K_2 . Further, facing a single equilibrium standard K , the measure of students from type u university is pinned down by optimization—each university chooses a grading standard that maximizes the expected product of those receiving good jobs. \square

Proof of Proposition 10. Suppose only students with “A” grades receive good jobs. Equilibrium requires $E_{HA}[\theta] = E_{IA}[\theta]$. Define $Q_U(x) = \frac{\int_x^{\bar{\theta}} t f_U(t) dt}{\int_x^{\bar{\theta}} f_U(t) dt}$ for $x \in [\underline{\theta}, \bar{\theta})$, with $Q(\bar{\theta}) = \bar{\theta}$, to be the expected productivity of “A” students given any standard x ; $Q_U(x)$ is trivially strictly increasing in x . By CFOSD, $Q_H(x) > Q_I(x)$ for all $x \in [\underline{\theta}, \bar{\theta})$, so the value y defined by $Q_H(y) = Q_I(x)$ is less than x . \square

Proof of Proposition 11. Because “A” students are hired before “B” students, the social planner gives “A” to all students who, in her opinion, should be employed on a good job, and the labor market assigns good jobs only to “A” students. The first-order conditions to the social planner’s problem are:

$$\begin{aligned} -\alpha \hat{\theta}_H f_H(\hat{\theta}_H) + \alpha \lambda f_H(\hat{\theta}_H) &= 0 \\ -(1 - \alpha) \hat{\theta}_I f_I(\hat{\theta}_I) + (1 - \alpha) \lambda f_I(\hat{\theta}_I) &= 0, \end{aligned}$$

where λ is the Lagrange multiplier for the capacity constraint of the social planner problem. By the full support assumption, the densities are positive, so the first-order conditions simplify to $\hat{\theta}_H^P = \lambda$ and $\hat{\theta}_I^P = \lambda$. \square

Proof of Proposition 12. Observe that $Q_H^{-1}(\bar{\theta}) = Q_I^{-1}(\bar{\theta}) = \bar{\theta}$, and CFOSD implies that $Q_H^{-1}(\bar{\theta} - \varepsilon) > Q_I^{-1}(\bar{\theta} - \varepsilon)$ for any ε positive, but sufficiently small. Thus, $Q'_H(\bar{\theta}) < Q'_I(\bar{\theta})$. By the implicit function theorem, $\hat{\theta}'_H(\hat{\theta}_I | \hat{\theta}_I = \bar{\theta}) = \frac{Q'_I(\bar{\theta})}{Q'_H(\bar{\theta})} > 1$. Continuity of $\hat{\theta}'_H(\hat{\theta}_I)$ ensures that $\hat{\theta}'_H(\hat{\theta}_I) > 1$ over some non-degenerate interval, $[\tilde{\theta}_I, \bar{\theta})$. The result follows. \square

Proof of Proposition 13. By Proposition 12, there is an interval $[\tilde{\theta}_I, 1]$ where $\hat{\theta}'_H(\hat{\theta}_I)$ is at least 1, so that $\hat{\theta}'_H(\hat{\theta}_I) = 1$ when $\hat{\theta}_I = \tilde{\theta}_I$. Thus, for some $\hat{\theta}_H = \hat{\theta}_H(\tilde{\theta}_I)$ and $\hat{\theta}_I = \tilde{\theta}_I$, $Q_H(\hat{\theta}_H) = Q_I(\hat{\theta}_I)$ and $Q'_H(\hat{\theta}_H) = Q'_I(\hat{\theta}_I)$. The second derivative of $\hat{\theta}_H(\hat{\theta}_I)$ is

$$\frac{\partial^2 \hat{\theta}_H}{\partial \hat{\theta}_I^2} = \frac{Q''_I(\hat{\theta}_I)Q'_H(\hat{\theta}_H) - Q'_I(\hat{\theta}_I)Q''_H(\hat{\theta}_H) \frac{\partial \hat{\theta}_H}{\partial \hat{\theta}_I}}{(Q'_H(\hat{\theta}_H))^2} = \frac{Q''_I(\hat{\theta}_I)(Q'_H(\hat{\theta}_H))^2 - Q''_H(\hat{\theta}_H)(Q'_I(\hat{\theta}_I))^2}{(Q'_H(\hat{\theta}_H))^3}.$$

Thus, $Q'_H(\hat{\theta}_H) = Q'_I(\hat{\theta}_I)$, $\hat{\theta}_H(\hat{\theta}_I)$ is concave if $Q''_I(\hat{\theta}_I) < Q''_H(\hat{\theta}_H)$. We now solve for the shapes of derivatives of $Q(\cdot|\cdot)$ when densities have linear right tails:

$$\begin{aligned} Q(t|a, b) &= \frac{2}{1-t} \frac{\frac{a}{2}(1-t^2) + \frac{b}{3}(1-t^3)}{b(1+t) + 2a}, \\ Q'(t|a, b) &= \frac{2(a+bt)(3a+2b+bt)}{3(2a+b+bt)^2} = \frac{2}{3} \left(1 - \frac{(a+b)^2}{(2a+b+bt)^2} \right), \\ Q''(t|a, b) &= \frac{4b(a+b)^2}{3(2a+b+bt)^3}. \end{aligned}$$

Observe that $Q'(\hat{\theta}_H|b_H) = Q'(\hat{\theta}_I|b_I)$ implies $\frac{a_H+b_H}{2a_H+b_H+b_H\theta_H} = \frac{a_I+b_I}{2a_I+b_I+b_I\theta_I}$, which in turn combined with $a_H b_I (1 + \hat{\theta}_I) < a_I b_H (1 + \hat{\theta}_H)$ implies $\frac{b_H}{b_H(1+\hat{\theta}_H)+2a_H} > \frac{b_I}{b_I(1+\hat{\theta}_I)+2a_H}$. Multiply both sides of this inequality of derivatives by $\frac{a_H+b_H}{b_H(1+\hat{\theta}_H)+2a_H}$ and $\frac{b_I+a_I}{b_I(1+\hat{\theta}_I)+2a_I}$, respectively, and multiply by $\frac{4}{3}$:

$$\underbrace{\frac{4b_H(a_H+b_H)^2}{3(2a_H+b_H+b_H\hat{\theta}_H)^3}}_{Q''(\hat{\theta}_H|a_H, b_H)} > \underbrace{\frac{4b_I(a_I+b_I)^2}{3(2a_I+b_I+b_I\hat{\theta}_I)^3}}_{Q''(\hat{\theta}_I|a_I, b_I)}.$$

Therefore, $\theta_H(\theta_I)$ is concave at $\tilde{\theta}$. Since $\frac{\partial \hat{\theta}_H}{\partial \hat{\theta}_I} > 1$ for every $\hat{\theta}_I > \tilde{\theta}_I$, concavity of $\theta_H(\theta_I)$ at $\tilde{\theta}$ contradicts $\theta'_H(\tilde{\theta}_I) = 1$. \square

Proof of Corollary 3. $-f_u(\theta_u)d\hat{\theta}_u$ is the increase in ‘‘A’’s at university u responding to $d\Gamma$. As increasing Γ causes grading standards to fall and $\frac{d\hat{\theta}_H}{d\hat{\theta}_I} > 1$, the result follows. \square

Proof of Proposition 14. University u solves:

$$\max_{\hat{\theta}_u, \hat{\mu}_u} \int_{\hat{\mu}_u}^{\bar{\mu}} \int_{\hat{\theta}_u}^{\bar{\theta}} \theta \mu dF_u(\theta) dG(\mu), \tag{C.1}$$

$$\text{s.t. } \hat{\mu}_u E_u[\theta | \theta > \hat{\theta}_u] = K. \tag{C.2}$$

The associated first-order conditions for interior solution are:

$$\begin{aligned} -g(\hat{\mu}_u)\hat{\mu}_u \int_{\hat{\theta}_u}^{\bar{\theta}} \theta dF_u(\theta) + \lambda E_u[\theta|\theta > \hat{\theta}_u] &= 0, \\ -f_u(\hat{\theta}_u)\hat{\theta} \int_{\hat{\mu}_u}^{\bar{\mu}} \mu dG(\mu) + \lambda \hat{\mu}_u \frac{f_u(\hat{\theta}_u)}{1 - F_u(\hat{\theta}_u)} \left(E_u[\theta|\theta > \hat{\theta}_u] - \hat{\theta}_u \right) &= 0. \end{aligned}$$

The densities are positive everywhere. Integrating and rearranging terms yields

$$E[\mu|\mu > \hat{\mu}_u]\hat{\theta}_u = \frac{g(\hat{\mu}_u)}{1 - G(\hat{\mu}_u)} \hat{\mu}_u^2 \left(\frac{K}{\hat{\mu}_u} - \hat{\theta}_u \right).$$

Solve for $\hat{\theta}_u$:

$$\hat{\theta}_u = \frac{K \frac{g(\hat{\mu}_u)}{1 - G(\hat{\mu}_u)} \hat{\mu}_u}{E[\mu|\mu > \hat{\mu}_u] + \frac{g(\hat{\mu}_u)}{1 - G(\hat{\mu}_u)} \hat{\mu}_u^2} \equiv R(\hat{\mu}_u). \quad (\text{C.3})$$

$R(\cdot)$ is an increasing function since $g(\mu) + \mu g'(\mu)$ is positive for all μ in the support:

$$\text{sign} \frac{\partial R(x)}{\partial x} = \text{sign} \left\{ \int_x^{\bar{\mu}} \mu \frac{g(\mu)}{g(x)} d\mu + g'(x)x \int_x^{\bar{\mu}} \mu \frac{g(\mu)}{(g(x))^2} d\mu \right\} = \text{sign} \{(xg(x))'\}.$$

Equilibrium is governed by equations (C.2) and (C.3). The latter is a decreasing curve in $(\hat{\mu}, \hat{\theta})$ space (see Figure 3.2) so there is a unique equilibrium described by their intersections. If the optimal solution is on the boundary with $\hat{\mu} = \underline{\mu}$, the optimality condition becomes $R(\underline{\mu}) \geq \hat{\theta}$; if it is at $\hat{\theta} = \underline{\theta}$, the optimality condition becomes $R(\hat{\mu}) \leq \underline{\theta}$.

The CFOSD assumption implies that $E_I[\theta|\theta > \hat{\theta}] < E_H[\theta|\theta > \hat{\theta}]$, for every $\hat{\theta} < \bar{\theta}$. Define $Q_u(\hat{\theta}_u) \equiv \frac{K}{E_u[\theta|\theta > \hat{\theta}_u]}$; then $Q_H(x) < Q_I(x), \forall x \in [\underline{\theta}, \bar{\theta}]$, and note that $Q_u(\hat{\theta}_u)$ is decreasing in $\hat{\theta}_u$. Then $\mu_{uA}^* = Q_u(R(\mu_{uA}^*))$ implies

$$\mu_{HA}^* = Q_H(R(\mu_{HA}^*)) < Q_I(R(\mu_{HA}^*)) < Q_I(R(\mu_{IA}^*)) = \mu_{IA}^*.$$

An analogous argument about $R(Q_H(\cdot))$ and $R(Q_I(\cdot))$ delivers $\theta_H^* < \theta_I^*$. \square

Proof of Proposition 15. The social planner chooses grading standards to maximize expected output subject to the constraint that job placement decisions are based on student social skills and the information contained in grades:

$$\max_{\hat{\theta}_H, \hat{\theta}_I, \hat{\mu}_H, \hat{\mu}_I} \alpha \int_{\hat{\mu}_{HA}}^{\bar{\mu}} \int_{\hat{\theta}_H}^{\bar{\theta}} (\mu\theta) dF_H(\theta) dG(\mu) + (1 - \alpha) \int_{\hat{\mu}_{IA}}^{\bar{\mu}} \int_{\hat{\theta}_I}^{\bar{\theta}} (\mu\theta) dF_I(\theta) dG(\mu)$$

$$\text{s.t. } \alpha \int_{\hat{\mu}_{HA}}^{\bar{\mu}} \int_{\hat{\theta}_H}^{\bar{\theta}} dF_H(\theta) dG(\mu) + (1 - \alpha) \int_{\hat{\mu}_{IA}}^{\bar{\mu}} \int_{\hat{\theta}_I}^{\bar{\theta}} dF_I(\theta) dG(\mu) = \Gamma.$$

The first-order conditions (assuming an interior solution) for a type u school are:

$$\begin{aligned} \alpha g(\hat{\mu}_{uA}) \hat{\mu}_{uA} \int_{\hat{\theta}_u}^{\bar{\theta}} \theta dF_u &= \alpha g(\hat{\mu}_{uA}) \lambda \int_{\hat{\theta}_u}^{\bar{\theta}} dF_u, \\ \alpha \left(\int_{\hat{\mu}_{uA}}^1 \mu dG \right) \hat{\theta}_u f_u \hat{\theta}_u &= \alpha \lambda \left(\int_{\hat{\mu}_{uA}}^1 dG \right) f_u(\hat{\theta}_u). \end{aligned}$$

The densities and α cancel on both sides. After canceling and isolating λ on the right-hand side, the equations for both types of universities become

$$\hat{\mu}_u E_u[\theta | \theta > \hat{\theta}_u] = \lambda, \quad (\text{C.4})$$

$$E[\mu | \mu > \hat{\mu}_u] \hat{\theta}_u = \lambda. \quad (\text{C.5})$$

Optimality condition, equation (C.5), is the same for both university types. If the optimal solution is on the boundary, $\hat{\mu}_u = \underline{\mu}$, the condition is $\hat{\theta}_u E[\mu] \geq \lambda$; if it is at $\hat{\theta}_u = \underline{\theta}$, the condition becomes $\underline{\theta} E[\mu | \mu < \hat{\mu}_u] \leq \lambda$. Denote the implicit function for $\hat{\theta}_u$ from equation (C.4) by $R_u(\hat{\mu})$ and that from equation (C.5) by $Q(\hat{\mu})$. Their derivatives are

$$R'_u(x) = -\frac{\lambda}{x^2} \overbrace{\frac{1}{(E_u[\theta | \theta > t])'_{t=\frac{\lambda}{x}}}}^{>1}, \quad Q'(x) = -\frac{\lambda}{(E[\mu | \mu > x])^2} \overbrace{(E[\mu | \mu > x])'_x}^{<1}.$$

Because both $F_u(\cdot)$ and $G(\cdot)$ feature increasing hazards, the derivatives of both expectations are less than 1 (Bagnoli and Bergstrom (2005)). As $E[\mu | \mu > x] > x$ for all $x < \bar{\mu}$, we have $0 > Q'(x) > R'_u(x)$. By CFOSD, the efficient job assignment equation (C.4) for a type H school is everywhere to the left of that for a type I school. Therefore, the intersection of equations (C.4) and (C.5) in $(\hat{\mu}, \hat{\theta})$ space that determine $(\hat{\theta}_H, \hat{\mu}_{HA})$ occur above and to the left of the intersection that determines $(\hat{\theta}_I, \hat{\mu}_{IA})$. Thus, $\hat{\theta}_H^P \geq \hat{\theta}_I^P$ and $\hat{\mu}_{HA}^P \leq \hat{\mu}_{IA}^P$, and $\hat{\mu}_{HA}^P = \hat{\mu}_{IA}^P$ occurs only when lower boundary on support of μ binds for both types of universities, so that all ‘‘A’’ students get good jobs.

References

- Aguiar, Mark, and Gita Gopinath.** 2006. "Defaultable Debt, Interest Rates and the Current Account." *Journal of International Economics*, 69(1): 64–83.
- Anson, Jack L., and Robert F. Marchesani.** 1991. *Baird's Manual of American College Fraternities*. Menasha, Wisconsin: Banta Publishing Company.
- Arellano, Cristina.** 2008. "Default Risk and Income Fluctuations in Emerging Economies." *American Economic Review*, 98(3): 690–712.
- Arteta, Carlos, and Galina Hale.** 2008. "Sovereign debt crises and credit to the private sector." *Journal of International Economics*, 74(1): 53 – 69.
- Austen-Smith, David, and Roland G. Fryer.** 2005. "An Economic Analysis of "Acting White"." *The Quarterly Journal of Economics*, 120(2): 551–583.
- Bagnoli, Mark, and Ted Bergstrom.** 2005. "Log-concave probability and its applications." *Economic Theory*, 26(2): 445–469.
- Bagues, Manuel, Mauro Sylos Labini, and Natalia Zinovyeva.** 2008. "Differential Grading Standards and University Funding: Evidence from Italy." *CESifo Economic Studies*, 54(2): 149–176.
- Bar, Talia, and Asaf Zussman.** 2011. "Partisan grading." forthcoming in *American Economic Journal: Applied Economics*.
- Beers, David T, and John Chambers.** 2003. "Sovereign Defaults: Heading Lower into 2004." Standard and Poor's RatingsDirect Research Report 09/2003.
- Betts, Julian R.** 1998. "The Impact of Educational Standards on the Level and Distribution of Earnings." *The American Economic Review*, 88(1): 266–275.
- Bi, Ran.** 2008. "Beneficial Delays in Debt Restructuring Negotiations." International Monetary Fund IMF Working Papers 08/38.
- Buchanan, James M.** 1965. "An Economic Theory of Clubs." *Economica*, 32: 1–14.
- Bulow, Jeremy, and Kenneth Rogoff.** 1989. "Sovereign Debt: Is to Forgive to Forget?" *American Economic Review*, 79(1): 43–50.
- Chari, V V, Patrick J Kehoe, and Ellen R McGrattan.** 2002. "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?" *Review of Economic Studies*, 69(3): 533–63.
- Chatterjee, Satyajit, and Burcu Eyigungor.** 2009. "Maturity, indebtedness, and default risk." Federal Reserve Bank of Philadelphia Working Papers 09-2.
- Chen, William, Li Hao, and Wing Suen.** 2007. "A Signaling Theory of Grade Inflation." *International Economic Review*, 48(3): 1065–1090.

- Coate, Stephen, and Glenn C. Loury.** 1993. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 83(5): 1220–1240.
- Costrell, Robert.** 1994. "A Simple Model of Educational Standards." *American Economic Review*, 84(4): 956–971.
- Das, Sanghamitra, Mark J. Roberts, and James R. Tybout.** 2001. "Market Entry Costs, Producer Heterogeneity, and Export Dynamics." NBER. working paper.
- De Paoli, Bianca, and Glenn Hoggarth.** 2006. "Costs of Sovereign Default." *Bank of England Quarterly Bulletin, Fall 2006*.
- Dubey, Pradeep, and John Geanakoplos.** 2010. "Grading Exams: 100, 99, ..., 1 Or A, B, C? Incentives In Games Of Status." *Games and Economic Behavior*, 69(1): 72–94.
- Eaton, Jonathan, and Mark Gersovitz.** 1981. "Debt with Potential Repudiation: Theoretical and Empirical Analysis." *Review of Economic Studies*, 48(2): 289 – 309.
- Fang, Hanming, and Peter Norman.** 2006. "Government-Mandated Discriminatory Policies: Theory And Evidence." *International Economic Review*, 47(2): 361–389.
- Hodrick, Robert J, and Edward C Prescott.** 1997. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking*, 29(1): 1–16.
- Hummels, David.** 2001. "Time as a trade barrier." Purdue University Research Report.
- Hummels, David, and Peter J. Klenow.** 2005. "The Variety and Quality of a Nation's Exports." *American Economic Review*, 95(3): 704–723.
- Livshits, Igor, James MacGee, and Michéle Tertilt.** 2007. "Consumer Bankruptcy: A Fresh Start." *American Economic Review*, 97(1): 402–418.
- Lizzeri, Alessandro.** 1999. "Information Revelation and Certification Intermediaries." *The RAND Journal of Economics*, 30(2): 214–231.
- MacLeod, W. Bentley, and Miguel Urquiola.** 2009. "Anti-Lemons: School Reputation and Educational Quality." National Bureau of Economic Research Working Paper 15112.
- Mailath, George J., Larry Samuelson, and Avner Shaked.** 2000. "Endogenous Inequality in Integrated Labor Markets with Two-Sided Search." *American Economic Review*, 90(1): 46–72.
- Milgrom, Paul R.** 1981. "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics*, 12(2): 380–391.
- Moro, Andrea, and Peter Norman.** 2004. "A general equilibrium model of statistical discrimination." *Journal of Economic Theory*, 114(1): 1–30.
- Norman, Peter.** 2003. "Statistical Discrimination and Efficiency." *Review of Economic Studies*, 70(3): 615–627.
- Ostrovsky, Michael, and Michael Schwarz.** 2010. "Information Disclosure and Unraveling in Matching Markets." *American Economic Journal: Microeconomics*, 2(2): 34–63.
- Rojstaczer, Stuart.** 2003. "Where All Grades Are Above Average." *Washington Post*.
- Rose, Andrew K.** 2005. "One Reason Countries Pay Their Debts: Renegotiation and International Trade." *Journal of Development Economics*, 77(1): 189 – 206.
- Rose, Heather, and Julian R. Betts.** 2004. "The Effect of High School Courses on Earnings." *Review of Economics and Statistics*, 86(2): 497–513.

- Ruhl, Kim J.** 2003. "Solving the Elasticity Puzzle in International Economics." University of Minnesota mimeo.
- Sandleris, Guido, Gaston R. Gelos, and Ratna Sahay.** 2011. "Sovereign Borrowing by Developing Countries: What Determines Market Access?" *Journal of International Economics*, 83(2): 243–254.
- Spence, Michael.** 1973. "Job Market Signaling." *The Quarterly Journal of Economics*, 87(3): 355–374.
- Todd, Walker F.** 1991. "A History of International Lending." In *Research in Financial Services*. Vol. 3, , ed. George Kaufman, 201–289. Greenwich, Conn.:JAI Press.
- Tomz, Michael.** 2007. *Reputation and International Cooperation: Sovereign Debt across Three Centuries*. Princeton, NJ:Princeton University Press.
- Tomz, Michael, and Mark L. J. Wright.** 2007. "Do Countries Default in Bad Times?" *Journal of the European Economic Association*, 5(2-3): 352–360.
- Wright, Mark L J.** 2001. "Limited commitment: Theory with applications to repudiation risk." PhD diss. University of Chicago.
- Yang, Huanxing, and Chun Seng Yip.** 2003. "An Economic Theory Of Grade Inflation." University of Pennsylvania.
- Yue, Vivian Z.** 2010. "Sovereign Default and Debt Renegotiation." *Journal of International Economics*, 80(2): 176–187.
- Zubrickas, Robertas.** 2010. "Optimal grading." Institute for Empirical Research in Economics - IEW.